

Burdens of proof of misconduct

Despite well publicized cases of scientific fraud, and less publicized examples of other misbehaviour by researchers, calls for enhanced external policing of science need to be greeted sceptically.

LAST week was a stimulating one for those who believe that the scientific process needs external policing. Martin Fleischmann and Stanley Pons, of "cold fusion" fame, lost a court case in Italy in which they had attempted to sue a journalist who linked their work too closely for their liking to with the concept of fraud (see page 369). A Brazilian virologist was successfully sued to the tune of \$50,000 for plagiarism of unpublished research (see page 371). And in the United Kingdom the *British Medical Journal* and *The Lancet*, acting in a combined initiative, called for the establishment of a central agency to tackle misconduct in research.

Nothing much is likely to happen fast in consequence of any of these. Most countries have been slow to follow the lead taken by the United States in establishing institutions to tackle this area. It is not surprising that the most vigorous proponents of such changes have come from medical research, given the potential dangers to patients. Such was the case in Denmark, the European leader in pursuing misconduct, whose medical research council played a key role in establishing the national Danish Committee on Scientific Dishonesty. This body includes scientists and lawyers nominated by interested parties that include, commendably, local political authorities. Other Scandinavian countries are actively pursuing similar possibilities, as is Australia. But Germany, France, Italy and the UK are prominent in their reluctance to respond to concerns in such an institutionalized fashion, to say nothing of Japan.

Is such reluctance advisable? That depends on whether those few cases of misconduct that are reported in such countries are merely the tip of an iceberg, or prove that, by and large, the science community's procedures — a combination of peer review, replication of results, employers' discipline and, occasionally, whistleblowing — are an effective form of self regulation. Certainly the factors such as citation records used increasingly to judge researchers' worth strengthen the pressure on individuals to behave badly — to bias their interpretation of data, or take too much credit for themselves. But evidence for an iceberg remains slender and anecdotal. That is one reason why calls for new agencies are unlikely to make much headway.

Significant expense is another. That is an obstacle which the United States Commission on Research Integrity (which advises Congress and the Department of Health and Human Services) may find insuperable. In its "Ryan report" produced late last year, the commission attracted the wrath of learned societies by proposing a range of initiatives that would impose additional costs on both the government and employers (as well as the societies). These included enhanced training and support for conferences in research ethics, and rigorously established procedures to deal with complaints, backed by quality control implemented with on-site visits by the Office of Research Integrity. (The report did not include an estimate of the costs.)

The critics are right to suggest that the Ryan report does not adequately make its case for such investment. But they would do well to express stronger concerns over the varied performance of institutions in carrying out their responsibilities, as reported by the commission. There is every reason to explore possible weaknesses in self regulation, in order to maintain public and political confidence in what is after all a costly public good.

The role of whistleblowers in particular can be critical. But it is also haphazard, and so, too, are the consequences of their actions, as evidenced by a recent ORI survey that documented a sizeable number of whistleblowers who experienced negative consequences. The Ryan report deserves support for promoting ways in which whistleblowers' interests can be more strongly protected.

European countries are right, given the apparently small scale of the problem, to be sceptical of the need for national agencies to address misconduct. But their institutions have tended to display too little energy in ensuring adequate self-policing. Stronger scrutiny, including systematic and independent investigation of researchers' experience of scientific bias and misrepresentation, is surely due. □

Avoid flaming rows

The future of the Internet as a research tool looks healthy, but there are dangers in excessively rapid communication.

"SPEAK in anger, and you will make the best speech that you'll ever regret", is a well-known maxim, as is one cure for curtailing this capacity — "count to ten". Many Internet users are now realizing that the same rules apply to comportment in cyberspace. The need to think twice when using the Internet has little to do with the predictable potholes of participating in 'flaming' (insulting other users) or the use of capitals in e-mail, which in Internet etiquette equates with SHOUTING. Rather, it has to do with the new problems created both by the speed with which information can be transmitted over the Internet or put on a Web page, and the mutation in mentality that some individuals undergo when they get behind the wheel of their computer.

This week's Briefing (pages 377–381) takes a serious look at the Internet and where it is going. But along the way *Nature* assimilated some anecdotal evidence of disadvantages dogging the medium (some not a million miles from home). Some individuals who would usually refrain from attacking colleagues seem to lose all sense of restraint when behind the screen. Combine this Jekyll-and-Hyde transformation with the immediacy and telegraphic style of e-mail and you have the makings of diplomatic mayhem.

The speed with which data can be put on the Web has also led to cases where it has had to be retracted because it had not been analysed properly. Similarly, researchers have hastily posted data on the Web before its commercial potential had been considered. Scientists are old hands at shooting themselves in the foot by publishing before patenting. The Web has made the trigger even itchier.

The lesson is that e-mails and Web submissions should be left to cool in a drawer (or its electronic equivalent, the draft folder) for a duration that is directly proportional to their scientific importance or potential for controversy. And even though its deployment can seem nauseatingly twee, scientific manners could only benefit from greater use of :-).



US physicists greet promise of open access to CERN's hadron collider

Washington & Munich. The European Laboratory for Particle Physics (CERN) in Geneva, Switzerland, and the US government have agreed to proceed with detailed negotiations on the participation of US scientists in the planned Large Hadron Collider (LHC) project on the basis of a US contribution of approximately \$530 million.

Officials continue to stress that the agreed figure is intended only as a basis for negotiation, and does not constitute a commitment of funds. But physicists on both sides of the Atlantic are jubilant at the news, which appears to open the way for the largest ever US involvement in a scientific facility built overseas.

The negotiations, which are due for completion by December, will proceed on the basis of contributions from the US Department of Energy (DOE) of about \$225 million for accelerator development and construction and about \$225 million for the particle detectors, in addition to \$80 million for the detectors from the National Science Foundation (NSF), according to a statement issued jointly last week by CERN and DOE.

The tentative agreement indicates that the two sides have overcome wide differences about the necessary scale of a US contribution (See *Nature* 379, 757; 1996). According to US officials, the DOE made its offer of \$450 million "plus or minus \$50 million" at talks in Washington DC with Hubert

Curien, the former French research minister who is president of the CERN Council, on 29 February. It was accepted at a meeting of the council two weeks later.

"We have reached a very significant midway point in the discussions with CERN," says Martha Krebs, director of the DOE's Office of Energy Research and leader of the US negotiating team. She adds that "the significant number is really the mid-point" — \$450 million — and that US scientists would get open access to the facility, with the thorny question of restricting access "not an issue" in the negotiations.

Christopher Llewellyn Smith, director general of CERN, says that he is "thrilled by the progress of the negotiations", as well as by what he describes as the "constructive attitude" of the DOE and NSF.

Sidney Drell, deputy director of the Stanford Linear Accelerator Center (SLAC) in California, and chair of an advisory panel of US physicists that two years ago recommended a DOE contribution of \$400 million to the LHC project, describes the news as "absolutely terrific".

Neal Lane, director of the NSF, says that the foundation's proposed contribution has still to be agreed by the its governing body, the National Science Board. It would also

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Now with added American acceleration? A computer graphic of the LHC in the existing LEP tunnel.

"take a big bite" out of NSF's annual \$54-million budget for high-energy physics. But "CERN is going to be the most important place for high-energy physics, and it is extremely important that US scientists participate," Lane says.

US officials say that the numbers should be regarded as a basis for two negotiating groups, one dealing with the accelerator and the other with the detectors, to proceed with discussing the possible options, not as a firm US commitment.

That caution is echoed by Robert Walker (Republican, Pennsylvania), chair of the Science Committee in the US House of Representatives. The committee's strong support will be necessary — but not sufficient — to ensure funds for the project. A statement from Walker said he was "very supportive of US participation" in the LHC. But he did not endorse the suggested level of expenditure.

The White House Office of Science and Technology Policy (OSTP) has so far declined to comment, referring enquiries to DOE. But OSTP is said to be supportive of the negotiations, and — according to congressional staff — the low profile being adopted by the White House could help to keep the project out of trouble in Congress.

The Drell panel said that money for participating in the LHC could be taken from within the existing, \$670-million high-energy physics budget at DOE after 1998, when major construction projects at SLAC and at Fermilab in Illinois are due to be completed.

The DOE has requested \$15 million for LHC work in its 1997 budget, an increase from the \$6 million approved for this year. The largest part of the US contribution would be spent between 1999 and 2005, when the project is due for completion.

Colin Macilwain & Alison Abbott

Scientists lose cold fusion libel case

Munich. Stanley Pons and Martin Fleischmann have lost an 8-billion-lire (US\$6.3-million) libel case that they brought two years ago against an Italian journalist who referred to their work on 'cold fusion' as fraudulent.

The two scientists, who claimed to have achieved fusion of deuterium atoms at low temperatures while working at the University of Utah in Salt Lake City in 1989, had objected to some colourful wording in a review in *La Repubblica* by Giovanni Pace of a book about scientific fraud by Axel Kahn, a prominent French geneticist, called *False Prophets*.

They sued the newspaper, and were joined in their legal action by three Italian scientists from the University of Milan, Giuliano Preparata, Tullio Bressani and Emilio Del Giudice. These were not mentioned by name, but believed that, as fellow advocates of

cold fusion, they could be identified from the review (see *Nature* 363, 107; 1993).

But the judge ruled that Pace's words "represented an expression of the right to report and criticize on the part of the journalist, and as such are not derogatory", and ordered the five scientists to pay costs of L28.5 million. After hearing scientific evidence from experts representing both sides, the judge concluded that the claim by Fleischmann and Pons to have achieved cold fusion remained an unproven hypothesis.

He said that Pace's criticisms were justified "on the basis of the existence of important opposition from the scientific community, not just against the theory of the research and the way the experiments were conducted, but also the way the data were divulged and the conclusions reached about the future direction of research".

A. A.

Slow release of data adds to BSE confusion

Paris. The focus of European governments' reactions to the 'mad cow' affair shifted from science to politics last week, fanned by a desire to restore consumer confidence in European beef. As it did so, however, their actions came under increasing fire from scientists who claimed that authorities were now over-reacting to data suggesting a possible link between Bovine Spongiform Encephalopathy (BSE) and Creutzfeldt-Jakob disease (CJD), a fatal neurodegenerative disorder in humans.

Much of this anger has been focused on the British government, and in particular on the fact that details of ten new and unusual cases CJD reported to its independent 13-member Spongiform Encephalopathy Advisory Committee (SEAC) were not made speedily available to the scientists responsible for advising other European governments.

Ironically, on the eve of the UK government's announcement two weeks ago, the Edinburgh group that investigated the cases was due to present an update on the epidemiology of CJD in the United Kingdom at an international conference on spongiform encephalopathies in Paris. But they were recalled by the UK government and did not present the results.

"We are asking ourselves if this was censorship [by the British government]", says Olivier Robain, an expert on human and animal pathology of prions at the Salpêtrière hospital in Paris, who points out that the conference schedule could have been rearranged to allow the Edinburgh team to speak before leaving.

But criticism is no longer restricted to the British actions. At the beginning of this week, many scientists were expressing

concern at plans being discussed in Brussels for a widescale culling of British cattle, a move which they claimed could not be justified by the current scientific evidence, but was motivated primarily by public fear.

The repercussions of the way that scientific advice has been handled in the current crisis are likely to reverberate for a long time. For example, one official says that, after the British government went public on the SEAC report, the BSE/CJD research community was "completely paralysed" by

interview that it was unusual for scientific problems of this importance to be handled in the absence of available data, and that it is "unacceptable to trigger reactions of panic" on presumptions that had not yet been demonstrated scientifically, adding: "We have no serious data yet that suggests an epidemic."

Lazar's comments are echoed by German researchers. Hans-Dieter Klenk, a virologist at Marburg University, says that Britain's failure to make the results available have fanned public hysteria. Georg Pauli, head of the department of virology at the Robert Koch Institute, which is responsible for providing advice to the German government on scientific matters, also expresses concern that the data were not released, arguing that this would have allowed more rational discussion, and independent assessment of the danger.

Meanwhile the European Union (EU) has set up a high-level scientific commission to recommend avenues of research into possible links between BSE and CJD, to review existing safeguards and decide whether others are needed. The commission will be chaired by Charles Weissmann, the director of the Institute of Molecular Biology in Zurich, but its composition and detailed remit will not be known until Weissmann meets Franz Fischler, the European commissioner for agriculture, later this week.

Weissmann defends the actions now being taken by European governments to reduce the risk of contamination of humans through British beef products. "Even if the data are not convincing or scientifically conclusive, there is no choice but to act as if they were," he says, adding that it would be "irresponsible not to consider [the new and unusual cases of CJD] as highly suggestive evidence on which one should act".

He admits that more could have been done previously, arguing that, to his knowledge, no experiments have been done on the feeding of infected material to monkeys, for example. But he warns against recriminations based on the wisdom of hindsight. "We can say that could or should have been done", argues Weissmann. "But if we had taken all the steps we are taking now, people would have said why are you causing billions of pounds of damage by acting on fragmentary or non-existent evidence".

Weissmann says the measures taken by the British government were probably sufficient in the circumstances. The government's error, he says, was to claim that there was no danger of disease passing from cattle to humans, and then to tell farmers, abattoir workers and others that they should be "very careful about how you handle the cattle and the meat".

Declan Butler

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Facing the music: British experts Ray Bradley (left), John Pattison (centre) and Kenneth Calman (right) in Brussels.

the lack of the original data. Meetings between French researchers and government officials last week were also hampered by the lack of hard data, according to one French scientist.

"There was a fundamental error of not respecting the rules of scientific communication," says one researcher. Others claim that the public panic could have been reduced if information about the 10 British cases of CJD had been better presented. "Fear feeds on ignorance," says one.

Critics of the way that findings were presented include Philippe Lazar, director general of the French biomedical research agency INSERM. He said in a newspaper

France closes risk assessment body

Paris. In a week that saw worldwide media coverage of the possible risk to humans of 'mad cow' disease, the French government ironically abolished its only independent commission for assessing technological risks, le Collège de la Prévention des Risques Technologiques (CPRT).

The government gave no official explanation of why it decided to abolish the commission, which is attached to the prime minister's office and has produced reports on industries such as nuclear power, biotechnology and chemicals. Indeed, its abolition was announced in a three-line paragraph deep within a decree on the French language.

Rumours of the CPRT's imminent demise had been circulating since the

beginning of the year (see *Nature* 379, 4; 1996). Its members suspended their work in February in protest at apparent uncertainty over the government's willingness to maintain an "independent institution" with the authority to issue opinions on its own initiative.

Jean-Jacques Salomon, the chairman of CPRT, says that the government's motivation for abolishing the commission cannot be to save money, as it runs on a relatively low budget. He argues that either the government believes that the French "do not face any technological risks", or it does not appreciate criticism, claiming that the abolition was prompted by bodies such as the Atomic Energy Commission which consider the commission "a thorn in the side". **D. B.**

Geology hits rich vein as South Africa boosts science spending

Cape Town. The South African government is to increase its spending on science by 9 per cent in the financial year 1996–97, in line with the overall increase in government spending, and two percentage points above inflation. In a novel move, 7 per cent of this money has been reserved for research projects aimed at prioritized objectives.

The total budgets for the seven research councils, announced last month as part of the government's general spending plans for next year, amount to R1,002 million (US\$250 million). Almost three-fifths will go to the two largest councils, the Council for Scientific and Industrial Research and the Agricultural Research Council, which both spend all of their funds on in-house research.

In terms of individual disciplines, geology will receive a major boost, reflecting the economic importance of the mining industry, with significant increases to the Councils for Geosciences (up 29.7 per cent to R60 million) and Minerals Technology (up 16.2 per cent to R73 million). The Medical Research Council will receive an increase of 15.5 per cent, although its total will still be only R58 million.

The main research funding agency, the Foundation for Research Development, will suffer a slight drop of one per cent in its money from the science vote, to R86 million. But this will be compensated for by an extra R16 million from separate Reconstruction and Development Programme (RDP) funds earmarked for centres of excellence at universities, technikons and museums.

The two national astronomical facilities, the South African Astronomical Observatory and the Hartebeeshoek Radio-Astronomical Observatory, have both been given major increases of 23.4 and 22 per cent respectively in recognition of their important research contributions. The National Accelerator Centre, which has come in for criticism in recent years over its high operating costs, will suffer a small decrease, although its budget, at R38 million, still represents more than twice that of the observatories combined.

In line with the government's desire to emphasize strategic objectives, 7 per cent of this year's budget was reallocated on the basis of the recommendations of a panel of representatives from the private sector, civil service departments, universities and trade unions. The panel was chaired by Roger Jardine, director-general of the Department of Arts, Culture, Science and Technology. Council presidents made presentations to the panel to justify their bids in terms of how their research activities coincided with RDP priorities.

Michael Cherry

Brazilian researcher protests against 'plagiarism' fine

São Paulo. A Brazilian researcher has been told by an appeals tribunal that he must pay a fine of R\$50,000 (US\$50,000) to a colleague for what she claims to be plagiarism of her work, which she had discussed widely at scientific conferences but only published belatedly.

Carlos Augusto Pereira, a senior researcher at the Instituto Butantan in São Paulo, was first accused in 1991 by Yeda Lopes Nogueira, a researcher at the Instituto Adolfo Lutz, of plagiarizing her research on the cytopathic effects of the rabies virus in a McCoy cell line.

After Nogueira's accusations had been upheld by a court in 1994, Pereira's lawyers filed a counterclaim in the São Paulo state appeals tribunal, arguing that what Nogueira called plagiarism is standard practice among scientists. But he lost again, with all three tribunal judges voting against him.

Pereira can still appeal to the supreme court, but the unanimous tribunal verdict makes it hard for him to prove his innocence. "This decision puts Brazil out of line with the rest of world," says João Luís Macedo, one of Pereira's lawyers. "If someone in China wants to carry out research on something discovered here, would he need a permit to do so?", he asks, claiming that such a requirement would paralyse science.

Nogueira, a former researcher at the São Paulo Instituto Pasteur, claims to be the first to have discovered the effects of the rabies virus in a McCoy cell line, announcing her achievement during a meeting in Rio de Janeiro in 1982. An abstract summarizing her discovery was published in the report of the meeting. Although she did not publish her results in a scientific journal, either in Brazil or abroad, she reported them to a number of scientific meetings. In 1989, however, she learnt that a team led by Pereira was working on the same subject.

Both researchers subsequently sent papers on the effects of rabies virus on McCoy cells to the *Journal of Virological Methods*. Nogueira's was rejected, and when she learned that Pereira's had been published, she took legal action. Her lawyer argued that Pereira's paper violates her intellectual property rights, as it does not mention her presentations at meetings. She accepts that scientists often investigate the work of others, including replicating original research. But, it is a "very different matter" if a scientist "usurps another's idea, and attributes the other's findings to himself".

In the paper's acknowledgements, Pereira's team "thanks Ms. I. Nogueira for helpful comments". His lawyers quote this as a proof Nogueira knew of his research, while Pereira points out that she even

lectured at Butantan about her own work.

But her lawyer describes this acknowledgement as "another proof of bad faith", pointing out that it misspells her name. "He repeated the experiments, but he already knew my results," says Nogueira. "He added a few data, but nothing important."

The first guilty verdict against Pereira had argued that his 'plagiarism' had caused Nogueira "material and moral damage", and the new one confirms both the plagiarism and its moral impact. Nogueira claims that her scientific career was damaged, as she needed the research results to complete a master's degree.

Pereira is the more experienced researcher, having worked in Paris and Strasbourg for ten years before coming to the Butantan in 1985. He says that he has never denied Nogueira's contribution to the field "Our work is different, the contents are different, and we used different viral strains," he says. "Our basic conclusion is the same because as the same biological phenomenon was being studied," says Pereira, adding that he did not quote Nogueira in his paper because she had not published any papers he could refer to, apart from short abstracts in conference proceedings.

Ricardo Bonalume Neto

New grants for young Russian scientists

Moscow. President Boris Yeltsin last week signed a decree establishing 100 grants, each worth 60 million roubles (US\$12,000), which he will award personally to talented young scientists. The decree also hands over to Russian colleges and universities federal property they have been leasing for the past 10 years or more, in addition to the land they previously owned.

The decree was announced in a speech to the University Rectors Congress in Moscow, in which Yeltsin described intellectuals as one of the "major forces" in the struggle for voters' hearts during the current presidential election campaign.

The decree also authorizes the cabinet to present the state Duma, Russia's lower parliament, with a bill fixing the pensions paid to college and university teachers at 80 per cent of their salary prior to retirement.

Earlier in the week, Yeltsin sent a message of congratulations to the Joint Institute for Nuclear Research (JINR) on the fortieth anniversary of its official foundation in 1956. "Russia is proud that this world science centre was built on its territory, and our state will always support it," wrote Yeltsin.

Carl Levitin

Diploma planned to revive science in UK schools

London. British post-16 secondary education is to undergo its biggest overhaul for 40 years following the government's decision to introduce a diploma similar to the French *baccalauréat*, while strengthening existing academic and vocational qualifications.

At the same time, the government has promised to review science and mathematics education, following a significant fall over the past decade in the proportion of 16- and 17-year-olds studying only mathematics and science. Fewer than 17 per cent of pupils aged 16 or more studied science and mathematics at Advanced (A) level last year, compared to 30 per cent in 1984.

There is also evidence of "insufficient" mathematical ability among science undergraduates. But the proportion of pupils achieving top grades has been steadily rising; 21 per cent of those taking physics A-level examinations last year achieved the highest grade, compared to 13.3 per cent in 1989.

The plans have been recommended to the government by Sir Ron Dearing, who chaired a panel that carried out a 12-month inquiry into British education. They have been widely supported by representatives of industry, such as the Engineering Council and the Chemical Industries Association, which have long warned of the threat to Britain's industrial competitiveness posed by a poorly trained workforce.

The present vocational courses will be made more rigorous with the introduction of external assessment. The Advanced General National Vocational Qualification, the highest vocational qualification, will be renamed the 'Applied A-level'. These courses will be joined by a broader National Diploma, combining academic and vocational study in an attempt to stem the high drop-out rate from these courses.

The plans have been welcomed by the educational and scientific communities, but for a slightly different reason from that voiced by employers, namely that the proposals stop short of interfering with A-level qualifications. These have formed the central plank of British further education for more than four decades and provide entry qualifications to university courses.

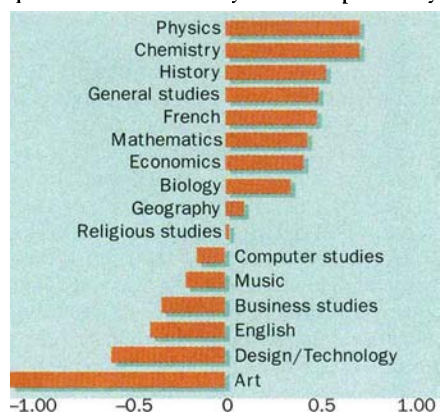
Although some people still consider A-levels to be the 'gold standard' of British education, others blame their relative narrowness — only two or three subjects are studied intensively over a two-year period — as a major cause of underachievement.

Concern that A-level standards may slip in the future has led to the government considering 'toughening-up' humanities subjects in line with science, mathematics and foreign languages, which are considered more 'difficult' (see chart, right). The plans also

include provision for the most able pupils to be "stretched" by studying degree-level courses while still at school.

"A-levels are a success story," says Carol Fitz-Gibbon, professor of education at the University of Durham, and author of a study suggesting that science is 'difficult' or 'severely graded' compared to other subjects. "They are a world class qualification. That's why we won so many Nobel prizes. But they are only for pupils at the top," she concedes. "Britain needs a vocational qualification that is equally well-respected."

Dick West, professor of education at the Open University, says the proposed new qualification is unlikely to be respected by



Degrees of difficulty: physics and chemistry were three-quarters of a grade more difficult than other subjects, according to an analysis of 1994 A-level results.

universities and employers while A-levels remain the 'gold standard'. "A-levels should have been abolished," he adds. West is among those who feel that the Dearing reforms "try to please everyone, but end up pleasing nobody".

West says he is unconvinced by Dearing's claim that fewer pupils study A-level mathematics and science because "society undervalues science". Mary Ratcliffe, chair elect of the Association for Science Education, says that poor science teaching, too, is one factor Dearing ought to have mentioned. The government, she says, has set a target of 1,000 science teachers to be trained by the end of this year. By February, only 86 places were filled.

Alan Smithers of Brunel University believes that the existing A-level curriculum works well for pupils intending to pursue careers in research. But he says it is of little practical relevance to the majority, who need to be "schooled in the nature and implications of science", but without the technical detail. "How relevant is an A-level in biology for a better understanding of the BSE crisis," he asks.

Ehsan Masood

Neutrino observatory can go ahead after Canada cancels cut

Ottawa. Intense lobbying by physicists in Canada and elsewhere has persuaded the Canadian federal government to reverse its decision, taken as part of a broad budget package last month, to end the support provided by Atomic Energy of Canada Limited (AECL) for the Sudbury Neutrino Observatory (SNO) Institute at the end of June (see *Nature* 380, 92; 1996).

Jon Gerrard, secretary of state for science, research and development, announced last week that AECL will continue to fund personnel attached to the project until construction is completed next January. Alternative sources of support will, however, still be needed to replace the individuals involved after the construction phase.

AECL has three scientists and several technicians working on the project, which has been planned to address the 'solar neutrino problem', and also provides some research facilities. These include neutron activation analysis used to detect small levels of uranium and thorium in construction materials.

Art McDonald, the director of the institute and professor of physics at Queen's University, stresses the important contribution of AECL's scientists to the project, whose funding comes from a broad spectrum of Canadian government and industry sources, as well as Britain's Particle Physics and Astronomy Research Council and the US Department of Energy.

If AECL had not been able to provide continued support, the project would have been in serious difficulties. One of those funded through AECL is the site project coordinator, Duncan Hepburn. "If we had lost him, we would have lost 50 per cent of our organization," says John Hykaway, a research associate in the physics department at Laurentian University. Hepburn has acted as a vital link between the project's many visiting physicists.

As to the international importance of the project, Hykaway reports that when the Nobel prize-winning physicist Hans Bethe recently celebrated his ninetieth birthday, a young scientist is said to have asked impertinently: "What are you hanging on for?" Bethe is said to have replied: "I'm waiting for the first data from the SNO."

According to Hykaway, the Sudbury project is the only one "that can really confirm whether neutrinos are flavour-changing", a process first invoked by Bethe and John Bahcall of Princeton University in New Jersey to explain the deficit, relative to theoretical expectations, of neutrinos that are observed to come from the Sun.

David Spurgeon

NASA: six weeks to save space station deal

Washington. The US Congress has given the National Aeronautics and Space Administration (NASA) six weeks to reach a watertight agreement with Russia that will put the troubled international space station back on schedule — and keep it there.

If NASA fails, say members of two congressional committees, the United States should abandon its collaboration with Russia and build the station only with its original partners, Japan, Canada and Europe.

The impasse results from the Russian government's failure to provide the funds needed to enable a contractor there to build the service module, the main Russian contribution to the station. US observers feel that Russia is dragging its feet in the hope of extracting more financial support from the United States.

"We may have to show that we are prepared to build it ourselves in order to get through to them and get them to stop testing us," Lori Garver, director of the National Space Society, said last week in testimony before a Senate subcommittee.

Al Gore, the US vice president, wrote to Victor Chernomyrdin, the prime minister of Russia, last month asking for assurances that funding would be made available. Gore asked for a response in time for last week's hearings. But nothing arrived, only adding fuel to congressional concern about the viability of the partnership.

At separate hearings last week of the space subcommittee of the House of Representatives Science Committee, and of the science, space, and technology subcommittee of the Senate Commerce, Science and Transportation Committee, Dan Goldin, the administrator of NASA, promised to resolve the problem within six weeks.

"If it goes beyond a month or so, we [will] begin to lose the ability to hold the schedule for the service module," he told the Senate subcommittee. "One month, maybe six weeks — no longer," he pledged.

Expanding on this before the House subcommittee two days later, Goldin said that the agreement would include "specific milestones" at three or four stages in the module programme to avoid a repetition of the situation. If an agreement was not forthcoming within six weeks, "we'll look at our contingencies and, at an appropriate time, act on them", said Goldin.

The Russian parliament has appropriated the money needed to build the service module, say Congressional staff. But this does not compel the finance ministry to release the money, and the contractor building the module is still not being paid. The issue is now being pursued by Thomas Pickering, the US ambassador in Moscow, and Goldin expressed confidence that it will be resolved.

But several members of Congress remain sceptical about any assurances provided by Russia. One is James Sensenbrenner (Republican, Wisconsin), chair of the House subcommittee, who went to Moscow in January and received what he considered a clear pledge from Oleg Soskovets, Russia's deputy prime minister, that the service module would be funded.

"That was January and this is March and they haven't done it," says Sensenbrenner. "I think I was misled by them. I feel very bitter about it — this administration is going to have to be much more hard-nosed."

Sensenbrenner is sufficiently upset to begin mentioning the unmentionable: a supplementary budget request this year, over and above the \$2.1 billion the space station

is already due to receive, to enable the United States to build the service module itself in time for its April 1998 launch.

It has been estimated that the United States is saving \$2 billion on the total cost of the space station through its partnership with Russia. But Sensenbrenner says that this number is inaccurate, and that the United States could drop its agreement with Russia at a much lower cost.

The module would cost \$500 million to build, and additional funds to launch. But Goldin says that the United States would be entitled to recover \$300 to \$400 million from Russia if it breaks off collaboration, sufficient to cover some of these costs.

NASA is likely to come under strong pressure to find any extra money from within its existing budget. But if this happened, as Sensenbrenner puts it, there would be "no science money left" at the space agency. Some Republicans would like some of the money to come from NASA's Mission to Planet Earth (see box, below).

Some of the criticism of the US-Russian collaboration is clearly political. Sensenbrenner, for example, has never liked the agreement, and Republicans are generally critical of what they see as the Clinton administration's "softness" in its dealings with Russia.

But concern goes much wider than party politics. Some supporters of the station would like the United States to break with Russia now, allowing the programme to be reconfigured in time for the planned commencement of hardware launches late next year. Such a move would also allow the current Congress, which may turn out to be more Republican — and therefore more supportive of the space station — than the next one, to appropriate the money needed.

Yet politics is likely to rule out such a clear-cut outcome. President Bill Clinton will probably await the results of the Russian presidential election in June, and the attitude of any new administration, before taking any precipitous action. Clinton himself will be keen to play down any problems with Russia in the run-up to his own election in November.

Congress sees the issue differently. Republican staff members say Russia has to understand that, if the service module runs late, it will be easier for Clinton to raise extra money from Congress to build it in the United States than to support a continued US-Russian partnership.

Meanwhile, opponents of the entire programme wait in the wings. "We've had eight designs already, and now it seems to me we're going to get a ninth," says Tim Roemer (Democrat, Indiana). "You will have some very tough questions to answer before the Congress if this agreement falls apart," he told Goldin. **Colin Macilwain**

Planet Earth funds revolve into 1996

Washington. The US National Aeronautics and Space Administration (NASA) has come under fire for carrying over \$650 million of the funding for its Mission to Planet Earth — one-third of the programme's annual budget — from the 1995 financial year (which ended on 30 September) to 1996.

James Sensenbrenner (Republican, Wisconsin), chair of the space subcommittee of the House of Representatives Science Committee, said last week that the amount of money carried over was "unheard of" and called it "a slush fund". He said that he would be asking the General Accounting Office, which reported the space agency's move, for a full and complete investigation.

According to Dan Goldin, the administrator of NASA, the money has built up

because managers have been reluctant to spend money in times of budget uncertainty. Goldin told Sensenbrenner at a Congressional hearing (see above) that he was "as concerned as you are" about the size of the carry-over.

The embarrassing accumulation of Mission to Planet Earth funds is likely to encourage critics of the programme, such as Robert Walker (Republican, Pennsylvania), chair of the House Science Committee, to push for even sharper cuts in its budget this year.

But the mission again received a ringing endorsement from Senator Larry Pressler (Republican, South Dakota), chair of the Senate Commerce, Science and Transportation Committee, who said that it was NASA's "most important and most relevant programme". **C. M.**

House backs curb on genetic information

Washington. The US House of Representatives last week passed a bill preventing health insurers from using genetic information to exclude potential clients, a move that many researchers say is needed to reassure volunteers for genetic screening programmes. The new prohibition is part of a large health reform bill designed to improve insurance portability.

But the Republican-authored bill includes many other controversial provisions that could prevent it from becoming law. Among these are plans to limit damages in medical malpractice suits, to allow individuals to set up tax-deductible 'medical savings accounts' to pay for routine health care, and to curb the regulation of insurance plans by state authorities.

The Senate has already drawn up a less ambitious bill, which has received broad bipartisan support. This has prompted Democrats to allege that Republicans in the House are merely trying to sabotage the more politically viable Senate legislation with their more radical version, approved on 28 March on a largely party-line vote.

Amid the larger controversy, the reference to genetic information was inserted in the House bill late last month "under the

very pleased," said Kathy Hudson, assistant director for policy at the National Center for Human Genome Research.

The insurance industry, however, claims that such a measure is unfair. In a voluntary marketplace, where many products are individually underwritten, companies "should have the same access to [genetic] information as the applicant", says Harvie Raymond, an assistant vice president at the Health Insurance Association of America.

Meanwhile, others are concerned that, even if controls on the use of genetic information are included in a bill that becomes law, insurers would remain free to demand

high prices for those who show genetic predisposition to disease. "That's a big loophole," says Neil Holtzman, director of genetics and public policy studies at the Johns Hopkins Medical Institutions in Baltimore, Maryland.

Holtzman chairs a task force on genetic testing of the NIH/Department of Energy Working Group on Ethical, Legal, and Social Implications of Human Genome Research (ELSI), which last month released draft principles endorsing a 1995 report from another ELSI working group that insurers be barred from using genetic information to discriminate. **Meredith Wadman**

Anthropologist cleared in patent dispute

San Francisco. The government of Papua New Guinea last week backed off allegations of wrongdoing against an anthropologist, Carol Jenkins, who has been at the centre of a row over the patenting by the US government of a cell line derived from the Hagahai people of Madang Province.

In doing so, it shelved any efforts to deport Jenkins, and agreed that her research had been conducted with the full consent of the Hagahai people, and that its benefits were intended to be shared among all concerned. At the same time, Jenkins has agreed to help government officials draw up a policy on the appropriate legal framework for research on indigenous populations.

Last month, Jenkins, a principal research fellow with the Papua New Guinea Institute of Medical Research who is involved in aid intervention research in the central highlands of Papua New Guinea, was escorted off a plane from Port Moresby that was heading to a meeting of the World Health Organization in El Salvador, and asked to explain her activities.

As she prepared last week to meet with Gabriel Dusava, Secretary for Foreign Affairs and Trade, it seemed that she might be deported. But after a long discussion, Dusava cleared her of any wrongdoing. In a subsequent press release, he emphasized that the meeting had resulted in an understanding of "the need for close government and research personnel co-operation in sensitive areas of research [such as] human blood and viruses, where formal frameworks are not available to set principles and regulations."

Jonathan Friedlaender, professor of biological anthropology at Temple University in Philadelphia, and a past director of the National Science Foundation's

physical anthropology programme, said that although the whole incident had been painful for Jenkins, it should help resolve conflicts over the social and ethical issues surrounding the patenting of modified DNA fragments taken from human subjects.

The US Patent Office issued a patent on the Papua New Guinea human t-lymphotropic virus (HTLV-I) to the Department of Health and Human Services in March 1995. It was the first human cell line taken from an indigenous population ever to be patented.

In an informal agreement, Jenkins, who as co-inventor was entitled to 50 percent of the royalties from any resulting commercial products, assigned her entire share to the Hagahai people. She says that she had discussed the matter with the Hagahai leaders, and that had jointly agreed that patenting would be the best approach.

At the time, scientists familiar with the research told newspapers that the likelihood of any commercial product being based on the cell line was slight, but that it might conceivably be used in a test for the variant virus, or to develop a vaccine. Nevertheless the patent claim outraged groups such as Rural Advancement Foundation International in Canada, a keen opponent of patents on living organisms — including humans.

Jean Christie, director of international liaison for RAFI, said the organization did not dispute Jenkins' intentions, but that major policy issues remained unaddressed. She said that Jenkins' agreement to work with the Papua New Guinea government was a positive step toward a major international debate over informed consent and patenting. But RAFI still insists that human genetic material and any other living organisms should not be patentable. **Sally Lehrman**



radar", says Lyle Dennis, director of the Genome Action Coalition, which is seeking a prohibition on the use of genetic information in both bills.

The House bill, like its Senate counterpart, would prevent insurers from denying coverage to employees when they change or lose jobs. It would guarantee continued coverage to those becoming ill. It states that insurers may not use 'health status' to bar people from coverage, and — unlike the Senate bill — defines this to include "genetic information."

"These two words make a good piece of legislation better," said Clifford Stearns (Republican, Florida), a key backer of the House provision, in a written statement. Those observing the issue from the National Institutes of Health (NIH) agree. "We are

Japanese genomics combines state and industry backing

Tokyo. Japan's genomics research received a major boost last week with the formal release of details of two new government-industry joint venture companies, the Helix Institute and Genox Research Institute Inc. (see *Nature* 378, 2; 1995).

Both will pursue the development and application of human genomic information. The Helix Institute will employ between 40 and 50 researchers and support staff, and will receive ¥6.6 billion (US\$60 million) over its planned six-year lifetime, about three-quarters from government and the rest from 10 private companies.

The new company's president is Teruhisa Noguchi, vice-president of Yamanouchi Pharmaceutical. Its director of research is Yukio Takigawa of Mitsubishi Chemical. According to Yoshiro Ohtaki of Japan Associated Finance Corporation — a venture capital fund participating in the project — Takigawa was chosen for his experience as head of planning at another government-industry 'bio-venture', the Protein Engineering Research Institute in Osaka.

The new institute's laboratories are to be located at the Kazusa DNA Institute, a molecular genetics research centre set up by the Chiba prefectural government in Sarazu, east of Tokyo.

Ohtaki says that staff for the new institute will be recruited from the contributing companies and from the academic community, both in Japan and overseas. Three expatriate Japanese researchers from the US biotechnology company DNX have already been recruited. The research aims include development of methods for identifying functional genes, high-efficiency cloning technology and cellular systems for expressing functional proteins.

Genox Research Institute Inc., which will mainly pursue medical applications of genomic information, has been established with support from a venture capital fund operated by the Ministry of Health and Welfare (MHW) and from eight private companies. The institute will receive ¥4 billion over its seven-year lifetime, half of which will come from the government.

Genox, which will be based at the National Centre for Children's Disease in Tokyo, plans to develop a system for cross-referencing genetic and clinical information, paying special attention to allergy. The institute's president, Haruo Naito, and head of research, Masaji Ono, are both executives of Eisai Pharmaceutical, one of the participating companies. Seven staff will be employed initially, but this may eventually increase to 20. Stephen Barker

Researchers rally to save primate pioneer's station

Giessen & London. A group of scientists and historians are stepping up an international effort to save the 80-year-old remains on Tenerife, in the Canary Islands, of one of the world's first research stations on primate behaviour, now in danger of being demolished to make way for a luxury housing complex.

More than 300 academics, including the directors of major primate research stations in Russia, Japan and the United States, have written letters and signed petitions urging the island authorities to preserve the anthropoid station, which made a name for its director, Wolfgang Köhler. His book *The Mentality of Apes*, one of his two major works that originated at the station, is considered a seminal contribution on chimpanzee intelligence.

"Köhler was one of the founders of the school of Gestalt psychology, and in effect a grandparent of cognitive psychology, its successor," says Robert Glaser of the University of Giessen in Germany who is coordinating the campaign to preserve the station.

Supporters of the campaign include Jane Goodall, scientific director of the Jane Goodall Institute in Hampshire, England, who says the building should be preserved as a museum. "Köhler contributed greatly to our understanding of our closest living relatives," wrote Goodall in a letter endorsing the petition. "His psychological tests still stand as examples of careful research that took account of the nature of the chimpanzee subjects," she says. "When I went into the field in 1960, *The Mentality of Apes* was my bible."

The station was established and sponsored by the Royal Prussian Academy of Sciences in 1913, and is located in the tourist resort of Puerto de la Cruz. It gave Köhler both the time and the space to observe reasoning in ten young chimpanzees that

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Near extinction? Houses could replace ape station.

had been brought from west Africa.

Known as La Costa because of its proximity to the coast, the research station contained a large experimental playground area, an ape house, a laboratory and Köhler's house, Casa Amarilla, where he lived with his wife, Thekla, and their young children. But the site now lies on land earmarked for development into luxury housing, despite being designated in 1994 as a site of historical interest by the island council of Tenerife, the Cabildo.

The station, in addition, has since been damaged twice. "The walls are crumbling, the house has no roof and is in danger of collapse," says Austin Baillon, a retired oil company executive and a spokesman for the Wolfgang Köhler Association. Baillon believes the damage could be the work of the owners of the land. "The roof was taken down in the hope that the building would collapse due to rain and not bulldozers," he claims.

The building's present owners, believed to be a company called Canary Property, have declined several requests for an interview. The land was owned from 1918 to 1994 by the Yeowards, a British family, once involved in merchant shipping, that has had a long association with Tenerife. Requests for interviews with family representatives were not answered.

Ehsan Masood

India signs up to CERN's new accelerator

New Delhi. India has agreed to make a contribution estimated to be worth \$25 million towards the construction of the 7-TeV Large Hadron Collider (LHC) planned for the European Laboratory for Particle Physics (CERN) in Geneva.

According to the terms of an agreement signed in New Delhi last week by R. Chidambaram, the secretary for the Indian Department of Atomic Energy (DAE), and Christopher Llewellyn Smith, the director general of CERN, India will supply 12,000 superconducting sextupole corrector magnets that will be located along the path of

the proton beam passing through the 27-kilometre circumference tunnel.

Indian scientists will also develop and supply software to control the synchrotron that will accelerate the protons to be injected into the LHC. A spokesman for DAE said that Indian scientists are also interested in participating in the construction of the two LHC detectors. He said that scientists from the Tata Institute of Fundamental Research in Bombay, DAE's own installations, and many universities will be involved in the construction of the detectors.

K. S. Jayaraman

Anticancer drugs to gain speedier approval in United States

Washington. President Bill Clinton last week announced a new policy intended to speed up the approval of anti-cancer drugs, as Republicans in Congress pushed ahead with legislation to expedite the US drug approval process.

Clinton announced that the Food and Drug Administration (FDA) will no longer require anticancer drugs to demonstrate increased patient survival time or quality of life before going to market. Instead, new drugs will need to show only tumour shrinkage for market approval. But companies will still have to complete studies showing long-term effectiveness once a drug is on the market.

The administration estimated that at least 100 of more than 200 anti-cancer drugs now being developed in the United States could have their approval times shortened under the new rules.

The FDA also announced that it will begin soliciting applications from makers of experimental cancer drugs that have been approved abroad, but not in the United States, for use in programmes that speed drugs to patients with life-threatening illnesses. □

Breast cancer test 'premature'

Washington. A private institute in Fairfax, Virginia, is offering a genetic test for a mutation of the *BRCA1* gene, which is linked to hereditary breast cancer. This is despite the view of leading geneticists that such tests are not yet ready for public use.

The Genetics and IVF Institute is selling the test, which detects a particular mutation known to be linked to breast cancer in Jewish women, for \$295. Geneticists involved in work on the so-called breast cancer gene had agreed not to make testing available until they knew more about how to interpret test results. □

Clearance sought for abortion drug

Washington. The Population Council, the group which holds US patent rights to the French abortion drug RU-486, has filed an application to the Food and Drug Administration (FDA) for approval of its use. The council also says that it has given manufacturing and distribution rights for the controversial drug to a new company, Advances in Health Technology.

The FDA, which normally takes 6 to 12 months to process such applications, has faced criticism from Republicans in Congress for the delay of its approval process. The agency can now expect to come under attack from them if it approves the use of RU-486. □

Agricultural scientists go on strike

Sydney. Staff at the Commonwealth Scientific and Industrial Research Organization's Division of Agricultural Research in Australia held a half-day strike last week in protest about expected job losses. The stoppage is only the third such strike in the organization's 70-year history. But the largely symbolic gesture may do little to stop one-third of the division's 280 staff and one of its two major sites closing down completely to bring costs back within the division's \$A13-million (US\$10-million) budget.

Oliver Mayo, the division chief, said he had considerable "sympathy" with the staff's position, but something needed to be done quickly to improve the division's poor financial state. Hard times in Australia's agriculture industry have caused the division's support from industry to dry up, while the government is keen to place greater emphasis on manufacturing research. □

Internet index of species launched

Manila. *Species 2000*, a worldwide initiative to create an Internet-based index of the world's known species, began last month in Manila, Philippines. The index will consist of a series of databases

covering each major group of organisms and will enable users to verify the scientific name, status and classification of every known species of plant, animal, fungus and micro-organism. The service will be made available as part of the Clearing House Mechanism under the United Nations Convention on Biological Diversity. □

Israel signs up to EU research

Brussels. Israel has become the first non-European country to participate in all non-nuclear research programmes funded by the European Union (EU). In return, European researchers will be given access to Israeli projects and results in reciprocal areas of cooperation. Under an agreement signed between Israel and the EU last week, Israeli research organizations and enterprises will participate in the 16 research programmes of the Fourth Framework Programme. Israel will contribute ECU30 million (US\$38.5 million) per year toward the Fourth Framework budget. □

UK research spending increases

London. Britain's total spending on research and development (R&D) increased by 3.7 per cent in real terms between 1993 and 1994 to a total of £14.6 billion, representing the steepest increase since the mid-1980s, according to figures released last week by the Central Statistical Office.

R&D in universities expanded particularly significantly, growing from £2.3 billion in 1993 to £2.6 billion. That carried out by private companies increased from £9.1 to £9.4 billion. Ironically, however, the strength of Britain's economy meant that expenditure on R&D as a proportion of gross domestic product fell from 2.20 to 2.19 per cent, a figure that compares to a peak of 2.29 per cent in 1986. □

Children think better to pop music

London. Listening to pop music appears to enhance reasoning skills more than other sounds such as Mozart and ordinary conversation, according to an experiment involving 11,000 children from 250 UK schools as part of *The Daily Telegraph's* Megalab experiment.

Groups of children were randomly divided into three groups: one listened to a Mozart quintet on the radio, another heard pop music and a third listened to speech. After a series of 'spatial reasoning tasks' at the end, the pop music group scored the highest marks.

The results may not be music to the ears of researchers at the University of California, Irvine, who have suggested that listening to Mozart improved spatial reasoning (see *Nature* 365, 611; 1993). □

Limits of lasers and stained glass

IMAGE
UNAMABLE
FOR A CORPUS
FOR A CORPUS
REASONS

Munich. Scientists and art historians restoring stained glass windows in churches in industrial east Germany have shown that newly developed laser techniques are less effective than traditional methods of hand restoration. Glasswork there has suffered from neglect, natural ageing and decades of severe air pollution. The picture above shows a nineteenth century window in Spremberg, Saxony, before (left) and after (right) restoration. □

Research and the Internet connection

Scientists are increasingly dependent on the Internet and the World Wide Web. Although global demand is currently causing congestion, deregulation and clever use of the Web point the way forward.

GAMMA-RAY bursts are one of astronomy's great mysteries. Occurring about once a day and lasting for about ten minutes, they are the brightest gamma-ray objects in the sky, and if they originate at cosmological distances — as some believe — they may be the most luminous sources in the Universe. But to understand the bursts, astrophysicists say they need to observe them at optical and radio wavelengths. They also need the Internet.

In a striking example of what can be done by exploiting the global reach of the Internet, two astrophysicists — Scott Barthelmy, from the Goddard Space Flight Center, and Gerald Fishman at NASA's Marshall Space Flight Center — have set up a fully automated rapid-reaction system linking an orbiting gamma-ray satellite with 30 optical and radio telescopes around the world. When a burst occurs and is detected on board the Compton Gamma Ray Observatory, the satellite alerts NASA's Goddard Space Flight Center in Maryland. Computers there quickly calculate the burst's rough coordinates and send them down the Internet. Seconds later, telescopes around the world turn towards the part of the sky where the burst is taking place.

Bursts are rare, brief occurrences, and precise coordinates are hard to establish. So far the telescope searches have failed to home in on their prey in time. But improvements to the targeting system are bringing the goal closer, says Fishman.

If astronomers are learning how important the Internet can be, other scientists are starting to see it as crucial in helping their discipline survive. Collaboration across the Net has suddenly become more important for the large fusion research centres in the United States, for example, following the massive staff cuts caused by the US Congress's decision earlier this year to cut funding for fusion research by a third. The Princeton Plasma Physics Laboratory (PPPL) says that the loss of one quarter of its staff means that collaboration with researchers at other universities over the Internet has become "essential".

PPPL recently began transmitting sound and video from the control room of its Tokamak Fusion Test Reactor over the Internet to collaborating scientists at other institutions. Meetings to plan experiments or to discuss results are also broadcast. During meetings, scientists can click on specially

prepared Web pages containing details of the graphics and texts being discussed. The scheme works well, in particular by allowing informal discussion before, during, and after experiments. "Researchers almost feel that they are here," says Steve Davis, head of systems development at PPPL.

International fusion research, too, may soon benefit from use of the Internet. A decision on whether to build the International Thermonuclear Experimental Reactor (ITER) must be taken before 1998. The four international partners in

Updating Web pages at CERN, the European Laboratory for Particle Physics where the World Wide Web was invented

ITER — Japan, Europe, Russia and the United States — are now studying the possibility of running ITER at a distance using the Internet.

Under the plan, the reactor would be operated 24 hours a day, with the control room rotating around the world from one participant to the next — scientists at the Lawrence Livermore Laboratory, in California, have already successfully tested the principle using the Tokamak Fusion Test Reactor at PPPL. Such a system could reduce the anticipated risk that when a site for the reactor is chosen, the other partners may lose interest.

Web comes to life

Oxford Molecular Group, a UK-based company producing drug-design software and services, has seen the writing on the wall for scientific software that runs only on PC, Unix, or Macintosh operating systems. The company is intending to shift its product range to run on the World Wide Web, using a Web browser as the standard interface. "Users are not going to buy products unless they are Web compatible," says Steve Gard-

ner, the company's senior product manager.

The decision reflects the potential of the Web, which runs on all computers and lets all users view information in exactly the same way. That power stems from the Web's origins — it was invented at CERN, the European Laboratory for Particle Physics, by Tim Berners-Lee, to let the various research groups in CERN's member states share information interactively.

Countless laboratories are now using the Web to help to manage collaborative research projects, while the world's major

public nucleotide and protein sequence databases have switched all their operations to the Web. Researchers can now simply use a Web browser to compare a newly isolated sequence with existing ones in the database, and to download any matches. By clicking on hypertext links in any one database, they can access any of the others.

Oxford Molecular's shift also reflects a move towards greater interactivity. At present, Web documents are fairly static objects. You can look at them, download them, and perhaps convert raw data into graphs by clicking on various plot options. But with most browsers that is about all. Indeed, the Web remains very similar to traditional publishing, with most users simply looking at information put there by others.

That may soon change. "Real money is now going into Web development," says Dave Raggett, from the World Wide Web Consortium (W3C), an industry organization set up by the Massachusetts Institute of Technology and the French computer research agency INRIA to develop the Web's potential, and which now has over 120 members, including companies such as Microsoft and Netscape.

One goal is to develop the browsers that allow Web users to create and modify documents on screen. Anything that could

David Parker/SPL

On other pages

■ Growth: boom or bust?	378
■ International packet jams	379
■ Late developers log on	379
■ End of the 'free lunch'	380
■ Cutting a long journey short	380
■ East Asia boosts bandwidth	381
■ High-speed testbeds	381

be done on a PC could also be done on the Web. Groups of users could work together on, say, a design of an aircraft. Authorized users would also be able to update experimental protocols or manuscripts.

JAVA, a programming language invented

Interactivity on the Web: Idfits Life, a Java application from Oxford Molecular, enables scientists to perform protein structure analysis over the Internet.

by Sun Microsystems, is also improving the ability of users to interact with Web pages. JAVA allows 'applets' to be written — purpose-built programs small enough to travel easily around the Internet. Applets for rotating molecules, or creating molecular models, can be stored on a Web server containing nucleotide or protein sequences, and called up by users, or even groups of users. Oxford Molecular and many other companies are designing their software to use JAVA. □

Internet growth — boom or bust?

"Scientific information should not have to compete on an equal footing with downloads of advertising icons," says Paul Ginsparg of the Los Alamos National Laboratory in New Mexico, and a pioneer of preprint distribution over the Internet. His comments echo the concern of many researchers that increasing business and public use is making the Internet congested and reducing its performance.

The Internet is now doubling in size every year. In 1983 just 200 computers were connected to it. Now the figure stands at almost 10 million, according to a survey published earlier this year by Network Wizards, a company based in California. And the real figure may be as high as 30 million, as many host computers are hidden behind the "firewalls" which protect private networks from being accessed from outside.

The Internet is growing so quickly because it provides a simple language — or Internet protocol (TCP/IP) — that lets any computer and communication systems talk to each other. The power of TCP/IP is that it allows networks to interconnect, and if you can connect to one you can connect to them all.

But the real boom has come from the World Wide Web, by far the most popular service on the Internet — it accounts for over half of all Internet traffic. Around 76,000 Web servers are connected to the Internet, according to Network Wizards, an increase of 2,400 per cent on last year.

The unanticipated increases in traffic are currently outstripping available capacity on many parts of the Internet, particularly on the transcontinental links (see page 379). Whether the Internet will be able to cope depends not on the number of users, but on how far increases in bandwidth — or telecommunications carrying capacity, expressed in bits per second (bps) — can keep up with the traffic that users generate.

Private networks appear to offer no escape. Even research networks that are cordoned off from the public Internet — called 'Intranets' — in countries such as France, are experiencing similar rates of

increase in traffic to research networks in the United States that share the public Internet. Moreover, private networks may be less well placed to install the extra bandwidth needed to absorb increases in traffic than countries that have integrated their research networks into the public Internet.

For one thing, research organizations that run their own networks have limited funds for buying new bandwidth in the form of dedicated cables. The French research network, Renater, can no longer afford to install bandwidth at the rate needed to keep

down prices. In Europe, bandwidth remains expensive because many telephone companies enjoy national monopolies.

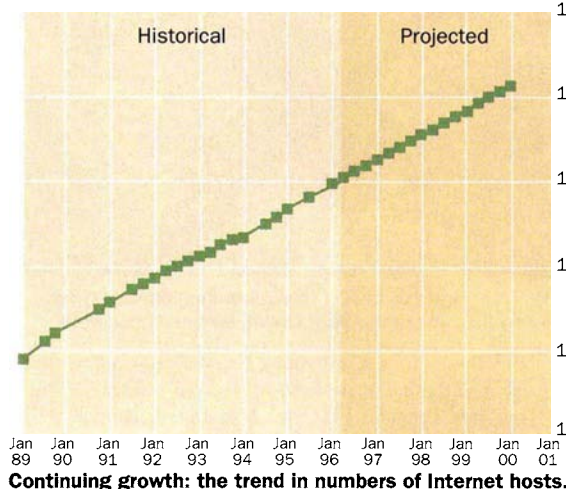
A leased 2 megabit per second (Mbps) line in Europe typically costs five to ten times as much as an equivalent line (1.5 Mbps) in the United States, says Brian Carpenter, group leader of communications systems at CERN. The exception is the United Kingdom, he says, where competition between British Telecom and Mercury has slashed the price of leased national bandwidth.

The difference between the United States and Europe is even more marked for high-speed links. It costs \$1 million a month to lease a 34 Mbps international high-speed link in Europe, says Rob Blokzijl, a nuclear physicist at the Science Park Watergraafsmere in Amsterdam, and chairman of RIPE, a body that coordinates European network service providers. In the United Kingdom, the research network JANET says that the £6 million it will spend this year to upgrade international links is "not nearly enough". If traffic growth continues, says JANET, the costs of upgrading will far outstrip available funds. Yet in the United States, the 45 Mbps link in Illinois between Fermilab and Argonne costs just \$50,000 a year.

Carpenter expects the situation to improve from 1998 onwards when Europe's national telephone monopolies must be privatized under a European Commission directive. "I am looking forward to deregulation," says Blokzijl. "Deregulate or die," says Carpenter.

One way to ease pressure on the networks is the growing use of regional caches (see box, page 380). But more draconian measures may be on their way. JANET is monitoring international traffic levels from connected research institutes and says that it is looking at the options for introducing controls.

Others have begun blocking the access of researchers to non work-related (often pornographic) sites. "This would be a last resort," says one official from the French research network, Renater. "We don't want to play gendarmes." □



pace with the 10 per cent monthly growth in traffic, according to Jean-Yves Babonneau, an official from the French computer research agency, INRIA. "We are completely saturated," he says.

Free markets required

John Quarterman, a US Internet analyst, predicts that bandwidth in the United States will be able to keep pace with traffic growth. In a competitive market, he argues, more users results in more income for Internet service providers who in turn will invest in more bandwidth to keep their customers happy. "The more users, the better it is for everybody," he says, pointing out the unit cost of bandwidth drops when you are able to buy it in large quantities.

Fierce competition among bandwidth providers in the United States has made bandwidth a commodity market, driving

Packet jams clog international highways

Bottlenecks at the ends of the Internet's international links are drastically reducing the networks' performance. "You can take

research organizations are also concerned that more international bandwidth might not increase performance. Domestic networks in some European countries run at slower speeds than the research network, so better international links might simply cause the bottleneck to migrate to the point of entry into a country.

Investment in better transatlantic links may similarly be a waste of money, complain French and UK researchers. Although average speeds within the US domestic network are good, the major gateway to the United States — "MAE East" in Washington DC — is becoming clogged as routers are overwhelmed by traffic loads. The United States urgently needs

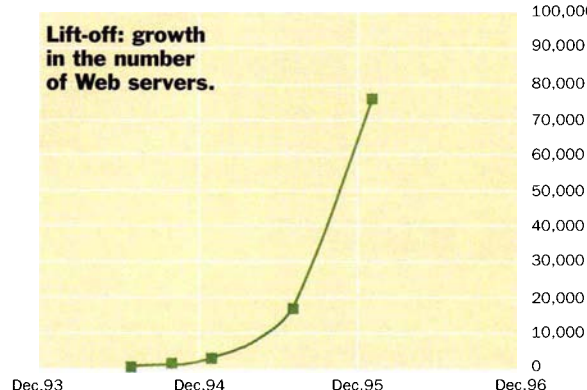
to install newer, bigger (and, of course, more expensive) buffers in these routers, they say.

Waiting for the Web

"Host contacted. Waiting for reply..." has become an all-too-familiar message to many users of the World Wide Web, and often it might as well read, "You have time to cook dinner". Heavy traffic on the Internet is one cause of poor performance. But many delays originate at the PC or workstation, and the computer it is communicating with. There are a host of limiting factors:

- Desktop connections: a 64 kbps link, for example — which transmits just 8 kilobytes of data per second (taking a byte as 8 bits) — will take an eighth of a second to send a 1-kilobyte e-mail, over 10 seconds to send an 85-kilobyte image, and nearly 3 hours to send a 10-second stretch of uncompressed video (around 75 megabytes of data). But the average desktop modem working at 14.4

Lift-off: growth in the number of Web servers.



Source: General Magic, Inc., February 1996

Concorde across the Atlantic, but you may still have to get the bloody underground from Heathrow to London," says Graham Cameron, head of services at the European Bioinformatics Institute in Cambridge.

To understand how Internet bottlenecks arise, one starts with the fact that the Internet transmits data as packets. Because packets are labelled with the address of their destination, they can travel independently through the network, taking many different routes. Packets from one computer travel down the local network where they mingle with packets from other users. At the junction between any network, for example a local and regional network, a self-contained device called a 'router' reads the address of each packet and forwards it to the next available router in this general direction. The process repeats itself until the packet reaches its destination.

When the traffic arriving at a link exceeds the link's bandwidth, routers store the excess packets in a buffer, waiting for a lull, as otherwise packet pile-ups would occur. This leads to delays, but under normal conditions still ensures that all packets reach their destination — the so-called 'best-effort' system. But if traffic is heavy, the buffers can become saturated. When this happens, the router simply dumps excess packets. The message fails to get through, and has to be sent again. The Internet's best effort is sometimes not quite good enough.

Bottlenecks and packet dumping are particularly common on international links, as packets compete for access to transcontinental or transoceanic cables. Traffic looking to use the link between the UK's JANET research network and the United States is often ten times the link's capacity, resulting in 90 per cent of it having to queue. Moreover, international traffic is growing faster than JANET can afford to install new bandwidth (see page 378).

Even if they could afford it, European

Late developers log on

Internet use in the less-developed regions of the world is booming. More than 30 countries connected to the Internet for the first time last year, including Albania, Bolivia, Cuba, Ethiopia and Uganda. The rate of connection in developing countries — although starting off from a lower level — often exceeds the overall growth rate of the Internet.

By the end of this year, only five or six African countries will lack Internet connections. Networks connecting research institutes and hospitals have been set up across the continent, in particular through cooperation with Orstom, the French research agency for development and cooperation, and Greennet, a non-governmental organization (NGO) that promotes environmental and human rights issues via the Internet.

In India, only a dozen or so universities are connected to the Internet, according to D. Balasubramanian, director of the Centre for Cellular and Molecular Biology in Hyderabad, though he expects this number to increase to between 80 and 100 this year.

Five years ago, Peru, Bolivia and the surrounding countries had no Internet connections. Now almost all universities are connected. In Peru, an Internet consortium of research centres, hospitals and NGOs has persuaded the government to support a 64-kbps satellite link that connects almost 300 institutions — and yet the country has only 3 telephones per 100 inhabitants.

Central America is becoming rapidly connected, according to Ann-Marie

Cetto, a theoretical physicist at the National University in Mexico City. The Mexican network covers the capital city, but also many universities in outlying regions. Universities and research centres are cooperating to provide links to smaller universities and countries in the region, she says, to reduce the new inequalities created between those who are connected and those who are not.

Cetto warns that Internet access could make developing countries even more dependent on research done elsewhere. "The Internet has made us sit up and think," she says, adding that the major need of developing countries is to reinforce their own research systems.

China and Russia have a problem in common with many countries more traditionally labelled as "developing" — the Internet's growth is restricted by poor telecommunications infrastructure. Earlier this year, the Chinese government also restricted international Internet access by domestic service providers to lines supplied by the ministry of post and telecommunications. The move was aimed at controlling the flow of information in and out of the country.

Russia has recently obtained access to full Internet services for the first time with the installation of the 110 Mbps Southern Moscow Backbone, linking Moscow and St Petersburg, that link in turn to other centres. The financier Georges Soros said last month that he would pay US\$100 million towards refurbishing the country's Internet infrastructure (see *Nature* **380**, 194; 1996). □

kbps will take four times as long to transmit the data.

- Poor protocol design: the same file, on the same server, at the same time of day can often be downloaded faster using File Transfer Protocol (ftp) than using the Web's HyperText Transfer Protocol (http). The current version of http also retrieves elements within a Web page (text, graphics, audio, animations) one after the other, so a page with 50 such 'objects' will require 50 back-and-forths (or 'flows') between the user and the Web server. This creates delays, because the duration of each flow

can be as long as that needed to transmit the data. The next version of http should retrieve several elements in a single flow.

- Icons also slow down Web servers: a block of server memory has to be set aside for several minutes to take account of the flow needed for each element embedded in a Web page — a precaution to catch late-arriving packets. The transfer of Web pages that are full of icons can quickly saturate the memory of even large servers.

- Underpowered servers: some servers are downright puny and cannot cope with their levels of access. Market forces should cause

operators to increase the performance of slow servers, as otherwise users will abandon their services.

- Impatient users: delays during Web sessions often indicate that Internet links are close to saturation. But instead of being patient, most people tend to call up the straggling Web page again by clicking 'Reload' in their browser. While this can speed up sessions for individuals, the overall effect of thousands of people simultaneously re-ordering Web pages is to send a new wave of packets through the Internet, pushing clogged links closer to breaking point. □

End of the 'free lunch' for big eaters

The Internet has never been free, but it usually feels that way at the point of use. Most Internet users now pay a flat monthly connection fee, after which they can use it as much as they like at no extra cost. So far, this system has sustained the Internet's exponential growth for over a decade, says Christian Huitema, from the French computer agency INRIA, and a former president of the Internet Architecture Board.

This pricing model is now being increasingly questioned. People who use up large amounts of bandwidth by running high-performance scientific applications, or sending bulky files of graphics, sound and video, will soon have to pay more than users sending small e-mail or text files, according to a growing number of Internet analysts, network specialists and computer companies.

Whether someone sends two or two thousand e-mails daily is no big deal in

terms of byte traffic. The Internet spreads out such difference among users by slowing down transmission when traffic is heavy and speeding it up when traffic is light. The Internet treats all packets alike.

But the system is now breaking down. A few users generating large amount of traffic can cause a disproportionate slowdown for others using the network. Frank Kelly, professor of the mathematics of systems at the University of Cambridge, says that pricing is needed to restrain such users and re-establish a fairer allocation of resources. "Otherwise the Internet will grind to a congesting halt," he predicts.

In some parts of the United States, applications such as remote control of instruments and supercomputing can no longer run properly over the Internet, according to Mark Luker, programme manager of the US National Science Foundation's very high

speed Backbone Network Service (vBNS). The problem lies in campus congestion as students flood to use the Web, he says.

To address the problem, NSF last month called for traffic from high-performance research applications either to be given priority on the Internet or to be coded so that it can be diverted to high-speed backbones (see *Nature* 380, 93; 1996). It also proposed that the Internet as a whole should also introduce prioritization of traffic.

The next version of the Internet Protocol — IPv6 or Ipng (next generation) — will allow streams of packets to be given priorities. Dave Ragget, of the World Wide Web Consortium, says he welcomes the NSF initiative, adding that he hopes it will catalyse the deployment of IPv6.

The pricing implications of prioritizing traffic are not yet clear. NSF wants a system in which users would pay first-, second- and third-class rates depending on the quality of service they want. Scientists using high-performance applications and multimedia users would pay first-class, for example.

This crude system may be sufficient to bring order to Internet traffic. At the same time, some scientists worry that it might prevent applications that need high bandwidth, such as videoconferencing, from becoming widely adopted by the scientific community.

What almost everyone agrees is that the NSF proposal is a better alternative than connect-time charges. These would inhibit individuals from exploring the Internet and dampen the vigour that made it a success in the first place, says Tony Rutkowski, vice president of Internet business development at General Magic, a Californian company. Usage-based charges would also cause unnecessary increases in the price of using the Internet, he adds, given that the expenses involved in billing can amount to over half the cost of the service provided.

Many researchers also oppose usage-based charges, predicting that their introduction would prompt laboratories to ration Internet use in much the same way as many now restrict access to international telephone lines. Budgeting for Internet use would also become much more difficult for research organizations. ►

Cutting a long journey short

Israel's seven universities have prohibited staff from directly accessing Web servers abroad, in a bid to cut the costs of leasing bandwidth on international links. Staff must now go via a device called a 'cache' that can cut international Web traffic by up to half by eliminating many duplicate requests for Web pages.

The principle of a cache is straightforward. If 1,000 people in a French town order Web pages from servers in the United States their requests will require 1,000 crossings of the Atlantic, as will the pages on the return journey. A cache is a local server which users would first check to see if the pages they want are already stored there. If so, the cache would supply them to the user. If not, it connects the user to the real server.

The cache would also make copies of these new pages for itself, which would then be available for future users, and so on. One can instruct a Web browser to automatically pass through a local cache by adding its address in the

Proxies' setting of the browser.

Caches are now being set up in many countries, including the United States, the United Kingdom, France and New Zealand, and they are becoming particularly popular in developing countries, which have difficulty affording international bandwidth. Caches can also save time and bandwidth within countries, in particular via a hierarchical system of interconnected caches established at the local, regional and national levels, says Hans Werner Braun, a staff scientist at the San Diego Supercomputing Centre who set up the NSF cache.

This can cause problems for time-sensitive information such as news. There are protocols being set up so that Web pages can define their own 'time-out' when they should be retrieved from source — but this can be abused. Caches also have knock-on effects, as page 'hits' — the number of times users access a page on a server — cannot be determined if they are being served from a cache. □

►“The research community will get the service that it wants within a highly competitive market,” says John Klensin, a senior data architect at MCI. Indeed, although bandwidth in Europe is likely to remain expensive until 1998, when national telephone monopolies will be dismantled, the prices of bandwidth in the United States have been reduced as a result of competition (see page 378).

Such competition can only intensify. The cable TV companies with their large optical fibre networks will soon overtake telephone companies as the major suppliers of bandwidth. Cable companies could afford to offer Internet connections simply as an extra on their regular services, argues Jon Crowcroft, professor of network systems at University College London, and a member of the Internet Architecture Board. And if the much-touted Information Superhighway can deliver megabyte-munching on-line video on an economic basis as it has promised to do, the costs of transmitting an e-mail or Web page would then be negligible. □

East Asia boosts bandwidth

Growth on Internet networks in Japan has lagged behind that of North America and Europe, but over the past few years the country has invested heavily in opening new Internet connections. The Internet has enjoyed strong support from several key members of the Liberal Democratic Party, the main party in Japan's coalition government, which allocated US\$100 million to various ministries and agencies to create Internet infrastructure.

Bandwidth on the link between SINET, the backbone connecting Japan's leading universities, and the United States was trebled to 6 Mbps in August last year. A month later, a 2-Mbps connection opened between Japan and Bangkok. Later this year, a 2-Mbps connection will open between Japan and Europe.

Within Japan, SINET's links have recently been upgraded to include 50-Mbps connection for nine national universities, running the length of the archipelago from Sapporo in the north to Fukuoka in the south. Another 20 universities and national research institutes operated by Monbusho — the education ministry, which runs SINET — are equipped with 6-Mbps connections. The same network links up a further 450 private and public universities with 1.5-Mbps connections.

Internet links between countries in the region are also improving. 1.5-Mbps links between Japan, Hong Kong and Singapore are expected to be completed later this year by the Asian Internet Holding company (AIH), a consortium of four companies from the three countries. □

High-speed testbeds herald new Net era

What might the next generation Internet look like? The Supercomputing 95 conference held in San Diego last December gives a glimpse into the future. Dozens of the US's supercomputers and state-of-the-art virtual environments were linked up for the first time for the event, and ran together fast enough to give the impression that they were all in the same room.

The I-WAY network, as it is called, uses clever engineering to link up a dozen high-speed networks, including NSF's new 155-Mbps national backbone, the very high-speed Backbone Network Service (vBNS). The idea is to bring the power of various US supercomputers to bear on individual problems, and provide computer systems of “unprecedented capability”, says Paul Woodward, director of the University of Minnesota's Laboratory for Computational Science and Engineering.

Woodward says that such systems would allow his team to make strong progress in work on geophysical and astrophysical turbulence. Supercomputer simulations of the merger of two vortices rotating at half the speed of sound (see photo, right) could be computed at different laboratories, for example, and then merged as the simulation progresses, he says. The process would also allow the simulation to be modified as it proceeds.

One of the first tests of I-WAY combined computing resources at the National Center for Supercomputing Applications (NCSA) and the San Diego Supercomputer to run simulations of Einstein's gravitational waves

and the evolving Universe. Combining resources in this way would also speed up research, says Woodward, by allowing exploratory problems to be carried out quickly, and even interactively.

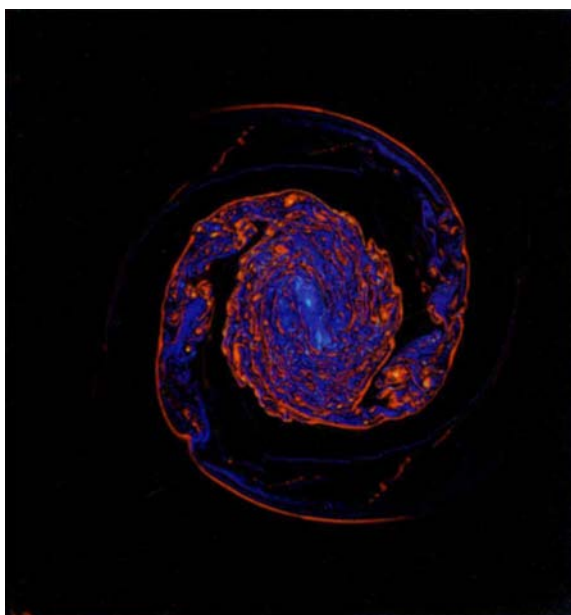
I-WAY was started by Tom DeFanti, from

tions of the next generation haven't been invented yet. We're trying to design and test the kind of network that will be there when they do arrive,” they say.

DeFanti, who carries out research on virtual reality and scientific visualization research, says that current visualization techniques are like “looking at a forest through a pair of binoculars; you might never realise there was a sky and a floor”. In the same way a drug binding site on a molecule might be “just up to your left”, he says, but might never be noticed using current techniques.

Virtual molecular modelling environments now under development go beyond static three-dimensional models, and immerse the scientists — or groups of scientists at different locations — in real-time molecular dynamics computer simulations combined with a virtual display of the chemical system. This lets scientists not only walk around the molecule, but also manipulate it by steering in real time a simulation run on a supercomputer. If the Internet evolves to high speeds as is expected (see *Nature* 380, 93; 1996), drug design using such techniques could soon be run from computers anywhere in the world. □

This Briefing was written by Declan Butler, *Nature's* European Correspondent, with additional reporting by Stephen Barker in Tokyo and Carl Levitin in Moscow. It is also available at <http://www.nature.com> and <http://www.america.nature.com> with hypertext links to many other sites.



Supercomputer simulations, such as this merger between vortices, will help understanding of astrophysical turbulence

the Electronic Visualization Laboratory at the University of Illinois at Chicago, along with Rick Stevens, from the Argonne National Laboratory in Illinois, and Larry Smarr, director of NCSA. Although I-WAY is aimed mainly at developing high-performance scientific computing, they believe that it will also unleash the next wave of evolution on the Internet. “The killer applica-

Implications of cloning

SIR — Our recent Letter describing the cloning of sheep by nuclear transfer from an established cell line¹ has created much popular interest about the biotechnological applications of the technique and the associated animal welfare issues. We wish to clarify these matters.

Five Welsh mountain lambs were born to Scottish blackface foster mothers. Gestation was extended by about 10 days. As reported, two of the lambs died within minutes of birth and a third after ten days. Independent post mortem analysis revealed incomplete urino-genital tract development in all three cases and of the vascular system in two, but in all other regards the animals were normal. The other two lambs are now 8 months old and healthy. The veterinary surgeon elected to deliver one of the lambs by caesarean section. It weighed 6.75 kg and died soon after birth. The others were born naturally without assistance and weighed 3.2, 4.15, 4.4, and 4.75 kg. The size of the lambs, with one exception, was within the range for both blackface and Welsh mountain ewes mated to rams of their own breed on the institute farm. The average birth weight for blackface lambs on that farm was 4.1 kg with a standard deviation of 0.61 kg and blackface ewes regularly give birth to lambs of over 6 kg following natural mating. During the past 4 years the largest two lambs weighed 6.6 kg. Welsh mountain lambs born on the same farm weighed between 1.9 and 4.9 kg. While he shared the obvious concern over the death of the lambs, the supervising veterinary surgeon identified no significant health or welfare problems in recipient ewes before or after lambing and although the long gestation periods caused initial concern, reported that they did not lead to significant problems at parturition.

A statistical analysis of the effect on birth weight of the procedures for nuclear transfer was not possible because there were so few observations and it was not clear what comparison should be made. We do not have birth weights of Welsh mountain lambs born to Scottish blackface ewes in our circumstances. However, it was to be expected that lamb birth weight would be greater than that of lambs born to commercial Welsh mountain ewes because the blackface foster mothers were of a larger breed and maternal size is well known to influence birth weight². Furthermore, all the lambs were single lambs, whereas commercial flocks produce a proportion of twins and even triplets, which are smaller at birth. In addition, the ewes were housed to allow careful treatment, rather than being run on the hills. All of these factors would be expected to increase the size of the lambs. Unusually large calves and lambs have been reported by others following a variety of techniques³, including

embryo culture either alone or in conjunction with pronuclear injection⁴ or nuclear transfer⁵. In the absence of appropriate controls, when it was not possible to know either the extent or the cause of the change, the single large lamb was not highlighted in our paper. The general response to manipulation or culture in these species is the subject of continuing research both in the United Kingdom and elsewhere.

The scientific importance of this work is that development was obtained from a cultured cell line. This technique will almost certainly be used in the first instance or biotechnological applications, such as the production of pharmaceutical proteins in milk, or the provision of organs for transplantation, rather than in commercial agriculture. Biotechnological applications are not dependent upon nuclear transfer for dissemination of transgenic stock, but rather the technique could be used to produce a small number of founder sires which would be used by traditional breeding.

An additional feature of this technique may be a reduction in the numbers of animals used in production of such transgenic founders. At present transgenic livestock can only be produced by injection of DNA into a nucleus in an early embryo⁶. This is inefficient and is associated with random integration, unpredictable expression, mosaicism, and variable germline transmission⁷. Typically several transgenic lines are generated for each application in order to be confident of having at least one with a high level of expression. By contrast, if it is possible to establish procedures for gene targeting in the cultured cells, then genes could be introduced specifically into active sites within the genome before nuclear transfer from cells of the desired genotype.

Although it may become possible to clone from adult cells, to date the technique has only been shown to work with a single embryo-derived cell type. Concern has also been raised at the possibility of cloning in humans. We do not envisage any situation in which, even if strictly possible, the application of similar techniques in humans would be either clinically useful or ethically acceptable. Such applications would also be illegal under existing UK legislation.

K. H. S. Campbell, J. McWhir
W. A. Ritchie, I. Wilmut
Roslin Institute, Roslin,
Midlothian EH25 9PS, UK

1. Campbell, K. H. S., McWhir, J., Ritchie, W. A. & Wilmut, I. *Nature* **380**, 64–66 (1996).
2. Walton, A. & Hammond, J. *Proc. R. Soc. B* **125** 311–335 (1938).
3. Walker, S. K., Hartwich, K. M., & Seamark, R. F. *Theriogenology* **45**, 111–120 (1992).
4. Walker, S. K., Heard, T. M. & Seamark, R. F. *Theriogenology* **37**, 111–126 (1996).
5. Wilson, J. M., et al *Anim. Reprod. Sci* **38**, 73–83 (1995).
6. Pursel, V.G., et al *Science* **244**, 1281–1288 (1989).
7. Carver, A.S., et al *Biotechnology* **11**, 1263–1270 (1993).

Imagine that

SIR — In his review of Nicholas Humphrey's *Soul-Searching — Human Nature and Supernatural Belief*, Tim Crane writes that it is a mistake to state that parapsychological phenomena cannot happen; it is enough to say they *do* not happen. "The fact that we can imagine extrasensory perception and the like is evidence for the fact that they are possible in some sense," he affirms (*Nature* **379**, 685; 1996).

Our imagination is evidence for nothing, except for the functioning, or dysfunctioning, of our brain. I am not sure in what "sense" impossible events or phenomena become possible. Throughout history men have imagined — and given very careful, meticulous and exact descriptions of — unicorns, dragons, giants, mermaids, gods, devils, angels, people whose heads grew out of their chests, heavens that revolve around the Earth and much else, to little or no purpose, so far as reality goes.

The fact that things do not happen may mean they are statistically very improbable but are not absolutely impossible, and our 'imagination' has nothing to do with it one way or the other. What has to do with it is physical law. It is absolutely impossible and not merely very unlikely that a heavier mass will orbit a lighter one. The 'fact' that we can imagine the Sun going round the Earth does not change matters and does not constitute evidence for such a thing.

If parapsychological phenomena are eventually proved to be real 'in some sense', that sense will be external physical law, and not because philosophers and other illogical persons have been imagining them.

Ralph Estling

The Old Parsonage,
Dowlish Wake, Ilminster,
Somerset TA19 0NY, UK

Willing genes

SIR — Daedalus suggests that narcotics users should receive the appropriate genes from the opium poppy or the coca bush which would allow them to synthesize their narcotics endogenously (*Nature* **380**, 206; 1996). There is a simpler way: we already possess the genes for making enkephalin, which has a similar effect. All that Daedalus has to invent is a way of turning them on at will.

M. F. Perutz

MRC Laboratory of Molecular Biology,
Hills Road,
Cambridge CB2 2QH, UK

Correspondence

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only, or e-mailed to nature@nature.com

Genetic testing and insurance

SIR — A thorough discussion of the implications of knowledge about genetic susceptibility to disease for health and life insurance (*Nature* 379, 389–392; 1996) is undoubtedly welcome. It is, however, important to clarify some of the issues raised about the difference between monogenic and multifactorial traits, between life and health insurance and also the distinction between population screening and the provision of testing services for at-risk individuals in the context of a clinical genetics service.

In the United Kingdom we have a National Health Service funded by taxation which provides for the treatment of diseases however severe and whether genetic or not; the genetic make-up of an individual is immaterial. Any change in the funding mechanism leading to the need to take genetic make-up into account could greatly damage prospects of using our growing genetic knowledge to improve medical care and develop targeted disease prevention programmes.

Life insurance is a more pressing issue. A screening programme should be carried out only if, as a result, some benefit can be recommended to high-risk individuals — ideally to reduce their risk to normal or near-normal. Where nothing can be done,

screening should be carried out only if individuals wish to know their genetic status, a wish they would very likely choose not to satisfy if they feared losing effective and appropriate life insurance.

It is important, as you emphasize, to distinguish monogenic from multifactorial inheritance. Only with monogenic inheritance is there likely to be a high risk of severe disease and so a significant issue for life insurance. Most of the monogenic diseases so far studied have a very early onset. When these are not treatable, they are of no relevance for life insurance.

For monogenic diseases with a later age of onset, such as most inherited cancer susceptibilities, there is often a chance of reverting the risk to near normal as, for example, with familial adenomatous polyposis, where the removal of all or most of the colon at an appropriate age results in a more or less normal life expectancy with a good quality of life. Here again, it is not clear that there is any relevance for life insurance.

The problems for life insurance, therefore, arise particularly when there are severe monogenic diseases of later onset for which nothing can yet be done, for example Huntington's disease, or for which treatment still leaves a shortened life

expectancy, as in cystic fibrosis. My own view is that it would be appropriate to accumulate case by case a specific list of diseases that are expected to raise a problem for life insurance, and to find a mechanism for them analogous to 'no fault' claims coverage.

For monogenic diseases, it is essential to distinguish between testing individuals at risk and large-scale population screening. In the family cancer clinic, a disease is known, or suspected, to be present in a family. Population screening is so far restricted to a small number of recessive diseases, especially phenylketonuria, which can be prevented when caught early.

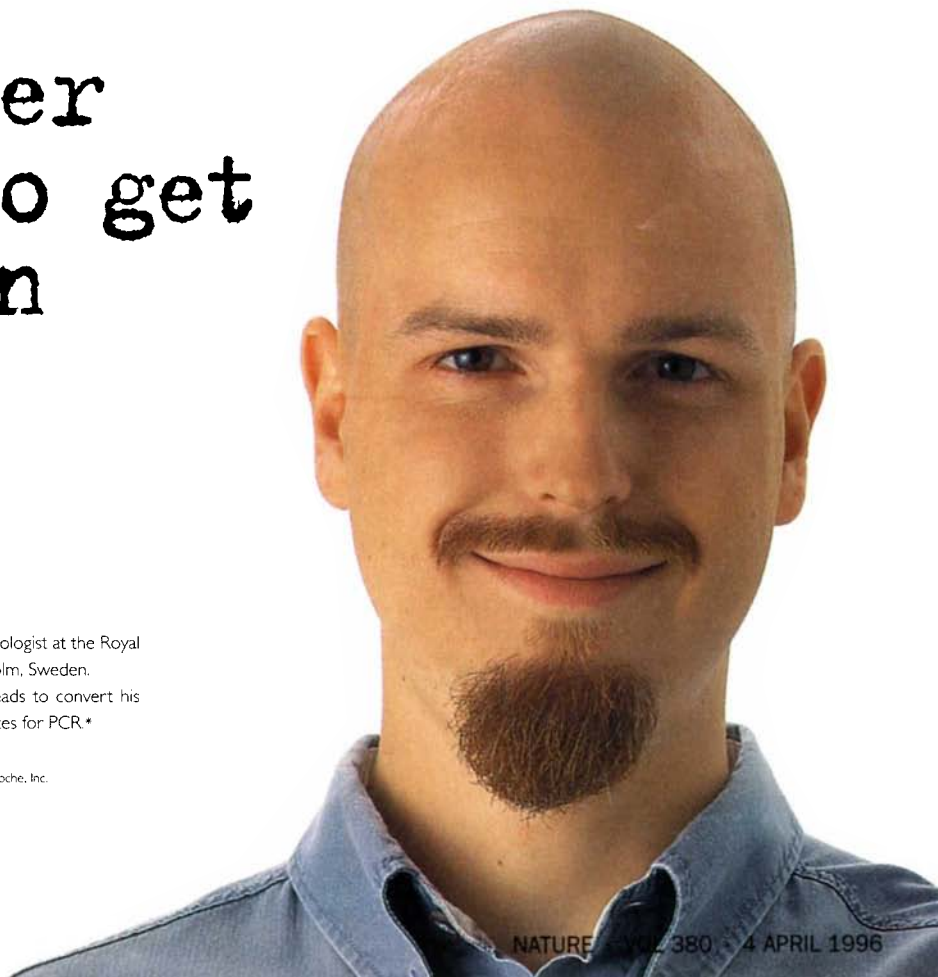
You rightly emphasize the difficulty of population screening for new mutations in complex genes such as *APC* (for familial adenomatous polyposis) or the recently discovered breast cancer genes *BRCA1* and *BRCA2*. But you do not distinguish clearly the differences between what is done in a family cancer clinic, where individuals followed up are known to be at extra risk, and what would be needed to justify population screening. The overall frequency of clearly inherited cancer susceptibility is probably less than 5% of the total. Population screening would obviously benefit sporadic individuals with new mutations and their subsequent families. The improved technology you discuss would be essential to such population

Patrik never fails to get a reaction

Patrik Samuelson is a molecular biologist at the Royal Institute of Technology in Stockholm, Sweden.

Patrik uses Ready-To-Go beads to convert his RNA samples into cDNA templates for PCR.*

* PCR is a patented process of Hoffmann-La Roche, Inc.



screening and, of course, ultimately and inevitably decisions will be made on some sort of cost-benefit analysis.

In multifactorial inheritance, there can be no certainty of disease and the probability of disease is determined by combining the effects of one or more genes together with environmental causes. The HLA determinant, *B27*, for example, is associated very strongly with ankylosing spondylitis, and nearly all individuals with the disease have *B27*. Nevertheless, only a small percentage of individuals with *B27* get the disease. The question then arises whether such information can be used to help the individuals at higher risk to overcome their susceptibility, perhaps by appropriate physiotherapy. For cancer, it might be early detection and for heart disease, it might be drug or dietary manipulation.

The aim is to have an effective form of genetically targeted disease prevention, whose end result is to achieve as near as possible an equalization of disease risks and effects between those with and those without a genetically determined susceptibility factor. In such cases, even more than for monogenic diseases, each situation will have to be carefully evaluated to justify population screening.

However, in such situations it is hard to see the relevance for life insurance: the screening programme would be carried out only if the risk of the genetically susceptible

could be brought back to near normal, and so the adjustment of a life insurance premium for a minor risk difference is unlikely to be worthwhile. Ultimately, if premiums could be adjusted exactly to an individual's risk, there would be no life assurance business, as this depends upon spreading risk.

I believe the problem of how to take into account genetically determined susceptibility to disease with respect to life insurance is manageable, while that for health insurance is fortunately not an issue in Britain, which still has a comprehensive National Health Service.

Walter Bodmer
(Director-General)
Imperial Cancer Research Fund,
PO Box No. 123,
44 Lincoln's Inn Fields,
London WC2A 3PX, UK

SIR — You highlight several of the vexing questions about the contending interests of those in need of protection — the public, and the life and health insurers. Serious debate about these issues is likely to become even more compelling and contentious as the extent of genetic testing increases in depth and becomes more sophisticated in predictive capabilities.

What especially caught our eyes was your suggestion of a programme of action that would permit both time for discussion and eventual confidence in the mechanisms

developed for judicious handling of genetic information. As you acknowledge, your proposal resembles the course followed in response to the controversy accompanying the introduction of recombinant DNA techniques.

Last fall, in a commentary on that period published in the *Proceedings of the National Academy of Sciences* (92, 9011–9013; 1995), we suggested that such a course of action could provide a model for dealing with a variety of difficult public policy issues, especially those associated with considerable scientific uncertainty. In particular, we pointed out that the model could be applied to the problem of suspected environmental hazards. Whether a moratorium, which was the first step in the response to the recombinant DNA issues and which you suggested as a first step in regulating the insurance companies' response to the genetic testing issue, is called for in each instance will be debatable. What is important is to establish guidelines or regulations that are initially deliberately rigorous, and a mechanism for prompt review and modification as knowledge and experience increase and misperception and prejudice diminish.

We foresee that the eventual emergence of new therapies for genetic diseases, now woefully lacking, will make early detection of impending disease or disability less problematic, even desirable. Moreover, as the ►

Do you always get a reaction when converting RNA into single-stranded cDNA templates for RT-PCR? Patrik does—thanks to a revolutionary polymeric bead that makes these critical conversions for him in new Ready-To-Go™ Kits.

The kits utilize the new Ready-To-Go bead that avoids many potential errors common to handling aqueous reagents.

That's because the "bead" is a complete,

pre-formulated, single-dose reaction—so there's little chance of your PCR templates becoming contaminated.

Just add your sample to the ambient-stable bead and watch it dissolve—that's all the pipetting required. In under a minute, you can begin full-length first-strand cDNA synthesis. This new approach to producing cDNA is only available from Pharmacia Biotech.

Just call us at 1 (800) 526 3593 in the United States, or +46 18 16 5011 from the rest of the world, for more information. Ask about the bead that never fails to give you a reaction.

 **Pharmacia
Biotech**
Uppsala, Sweden. (And the rest of the world)

number of genetic markers associated with disease increases, it is likely to become apparent that few of us are free from mutant genes that predispose or contribute to diseases of one sort or another. Who then will form the pool of risk-free, insurable persons? The strategy you propose would provide time to develop a constructive and humane response that can ensure access to appropriate care and security for all.

Paul Berg

*Stanford University School of
Medicine & Beckman Center for
Molecular and Genetic Medicine,
Stanford, California 94305, USA*

Maxine F. Singer

(President)

*Carnegie Institution of Washington,
1530 P Street NW,
Washington, DC 20005, USA*

SIR — Your Briefing failed to point out some of the underlying economic factors that do and will continue to influence why and how genetic tests are used. It also stops short of addressing the fundamental problems caused by the varied and mixed policies governments have adopted regarding genetic tests and insurance.

The benefits of these tests in insurance markets go not to the insurance companies but rather to the consumers of insurance with negative test outcomes and thus lower risk. In a competitive marketplace — and the insurance market is certainly that — these consumers will pay less for their medical and life insurance. Those who test positive will pay more. The insurance companies will ultimately be no better or worse off than they are now. The only difference is that they will get lower revenues from one group and higher revenues from another. Their profits should stay close to what they are now.

The real problem for the companies is in not having access to the same information as other companies or not being able to use it when others can. If insurers in one country can gather and use genetic information, they will create two policies where formerly there was one; one for those who test positive and one for those who test negative. The price to the latter will be less than they paid for the single policy offered before. These lower prices will attract the lower risk customers from countries not allowing the insurers the use of the information. Similarly, the prices of policies of those testing positive will rise and those individuals will seek insurance companies in countries that do not allow the use of genetic information. Companies unable to use the testing information will lose market share and be forced to raise their prices.

If this market switching were to continue, we would ultimately end up with only one insurance policy in each country; one for those who test positive in the country with no access to testing information and one for those who test negative in the coun-

try that allows the use of the information. The prices will be the same as they would be if both countries allowed the use of information and sold the two policies in their own countries.

Clearly, the use of testing information will benefit those with lower risks. Competition will provide them with lower prices. The poor people in that group will have a good argument for demanding they should not have to subsidize those with high risks. The group with positive tests will of course pay higher prices. What is important is that this will happen even if in some jurisdictions companies are allowed to use the information and in others not. The difference is that, in the latter case, large interim costs will be borne by both consumers and insurers as movement across the markets takes place. Unless a uniform policy is devised, we may well see this scenario tested. One could not design a better testing ground than a European Commission with free trade in insurance and mixed policies regarding the use of genetic testing.

Peter H. Nickerson

*Department of Economics and Finance,
Seattle University,
Seattle, Washington 98101, USA*

Science in India

SIR — John Maddox's leading article on "The prevalent distrust of science" (*Nature* 378, 435–437; 1995) and the correspondence that has followed (*Nature* 379, 292; 1996) make interesting reading here in India. Certainly, the funding patterns for research and teaching the world over seem to indicate anything but distrust for science. Moreover, in India, as in most postcolonial societies, science has consistently been projected by the state as the solution to all of society's ills: witness Jawaharlal Nehru's exhortation to his people to develop a "scientific temper".

But the institutionalization of science, which makes possible such 'icons' of development as hydroelectric dams, rockets and computers also ensures a widespread awareness that being a scientist is "just another job". And for most Indian scientists, the practice of science is largely a vocation that involves the learning, practice and teaching of certain techniques that bear little relevance to the conduct of politics, or, for example, to the marriage of one's daughter. In these matters, other 'unscientific' considerations often apply.

These compartmentalizations attenuate science's claims to social transformation, and lead to (a quite proper) distrust of those who speak on its behalf. This is quite separate from the distrust arising from "science in the service of death and destruction" that Stephen Keast mentions (*Nature* 379, 292; 1996).

This limited effectiveness is not a feature of the practice of science alone. The fact

that history and philosophy as disciplines are similarly institutionally 'contained' leads one to speculate that the relativization of knowledge systems is one strategy by which India 'manages' the potential for disruptive change; that such "distrust" is an important component of India's response to all-embracing systems.

A. Giridhar Rao

*Information Management
and Exchange Program,
International Crops Research
Institute for the Semi-Arid Tropics,
Patancheru 502 324, AP, India*

The last word?

SIR — There is a need for a word in taxonomy, and in medical genealogical, scientific, biological and other literature, that does not occur in the English or any other language. We need a word to designate the last person, animal or other species in his/her/its lineage.

An orphan is someone, usually a child, with no living parents. A foundling is someone, usually an abandoned baby, with no known parents. We do not have one word to describe the last person surviving or deceased in a family line, or the last survivor of a species.

Correspondence with etymologists and publishers of dictionaries to find a single word for 'the last of the line' in any language have been fruitless, with no word known.

A patient recently said: "I am the last of my line". His one word suggestion for his state was "omega", but several people known to us are named Omega and there are too many uses of this word for other purposes.

'Endling' was suggested when we were playing with possible new words. People active in the geriatric field thought it suitable too. Other suggestions were: lastoline (contraction for last of the line), yatim (Arabic for orphan or unique of its kind), Ender (one who ends or finishes), endmost (nearest to the end), endler (also a new word but *-ler* is not a recognized suffix).

Etymologists will recognize the two components for the derivation of 'endling' *end-* has several meanings, including 'extinction' and 'finish, concluding part'; *-ling* is a suffix added to denote 'connected with the primary noun' but also includes line and lineage.

We therefore propose that 'endling' be adopted to designate a person or one of a species that is the last of a lineage in his/her/its line. We are already using it when appropriate.

Robert M. Webster

Bruce Erickson

*Christian City Convalescent Center,
7300 Lester Road,
Union City,
Georgia 30291, USA*

Ownership of the human genome

S. M. Thomas, A. R. W. Davies, N. J. Birtwistle, S. M. Crowther and J. F. Burke

Despite wide interest in human genome patents, there has been little information about exactly who owns what. Europe would do well to ponder over the stake already held by Japan and the United States.

THE patenting of human DNA sequences is the most controversial aspect of the Human Genome Project (HGP). Academic scientists strive to increase knowledge and to publish their findings. Companies, on the other hand, regard the exclusivity gained through patenting as essential to successful commercial exploitation of the HGP for new therapeutic and diagnostic tools. Public debate over the ethics of patenting DNA has become more and more intense over the past three years as the HGP has gathered pace. Many religious and environmental groups, as well as individual scientists, remain fundamentally opposed to the granting of patents^{1,2}. Despite wide interest in the subject and its importance for public policy, there has been a remarkable dearth of even basic information about the actual ownership of human DNA patents. Instead, attention has generally focused on a few specific and often controversial examples of gene patenting as, for example, in the case of genes involved in breast cancer (*BRCA1*, *BRCA2*), cystic fibrosis and obesity^{3,4}.

Here we comment on the results of a comprehensive analysis of patented human DNA sequences using the GENE-SEQ database (Derwent Ltd, UK, release no. 18). Our data reveal that between 1981 and 1995 a total of 1,175 patents for human DNA sequences were granted worldwide. Each patent in our study contains an average of only three DNA sequences. The legal interpretation of the patent determines whether it protects only the sequence(s) disclosed, or whether its scope covers all possible forms of the gene (an

extreme example being Genentech's patent on tissue plasminogen activator).

Ownership is overwhelmingly dominated by the private sector (Fig. 1). Not only have three-quarters of these patents been granted to industry, but most of the 213 companies involved are Japanese or US. Nearly half these privately held patents are Japanese-owned, the others being divided almost equally between companies from the United States and the rest of the world. By contrast, the public sector is markedly under-represented. Only 17 per cent of the total have been granted to public institutions, and most of these are US.

Only the European Patent Office (EPO), the US Patent Office (USPO) and the Japanese Patent Office have issued significant numbers of patents. That half the world total is from the EPO is largely because both the United States and Japan have been vigorously patenting human DNA in Europe. Together these two countries own about 70 per cent of all EPO patents for human DNA sequences. Japan owns a third and the US nearly 40 per cent. Europe, on the other hand, has performed poorly, with only 24 per cent. Industrial investment has done much to boost the EPO total: more than 80 per cent are company-owned. Given that the United States has generally taken the lead in human genomics, it is surprising that only 16 per cent of human DNA sequence patents have been issued by the USPO. The paucity probably results from a backlog of biotechnology patent applications — a feature of the USPO over the past decade⁵. As

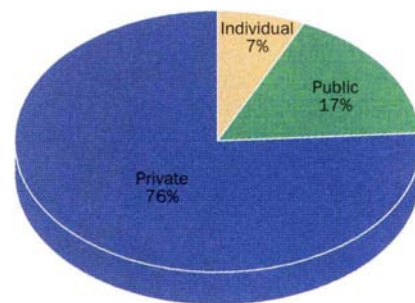


FIG. 1 Ownership of human DNA sequence patents. Includes all patents worldwide.

the situation improves, the number of USPO patents issued for human DNA sequences as a proportion of the world total is likely to increase because the United States is still viewed as the most important technological market for major inventions⁶. Only 41 per cent of USPO patents are privately owned and nearly all these belong to US companies.

The innovative character of small US firms dedicated to biotechnology is clearly reflected in our analysis (see table). Of the large US pharmaceutical companies, only three hold more than ten patents. In Europe, the situation is almost completely reversed. Large companies have been fairly active whereas the small-firm sector holds almost no patents. Remarkably, 88 Japanese companies own EPO patents for human DNA sequences. Although Japanese patents are generally narrower in their claims than USPO or EPO patents, the number of EPO patents held by companies such as Takeda is high at 44 (out of its world total of 63). Nevertheless, the Japanese consider their genome research to be lagging behind that of Europe and the United States by some five years.

Public-sector institutions in the United States own more than half of all USPO human DNA sequence patents and 59 per cent of publicly owned EPO patents. This is most probably due to the different institutional cultures of the United States, Europe and Japan. In the United States, public-sector scientists have for several years been encouraged to assess the patentability of their research results. Well-developed collaboration between US universities and industry and an environment that fosters academic entrepreneurs have undoubtedly stimulated patenting in the public sector. Why have European

PATENT OWNERSHIP OF PUBLIC- AND PRIVATE-SECTOR ORGANIZATIONS

Private company	Country of origin	No. patents	Public institution	Country of origin	No. patents
Takeda	J	63	US Department of Health and Human Service	US	28
Genentech	US	41	University of California	US	13
Immunex	US	23	National Institutes of Health	US	11
Teijin	J	23	University of Washington	US	11
Hoffmann-La Roche	S	18	University of Texas	US	7
Ciba-Geigy	S	16	Japanese Foundation for Cancer Research	J	6
Suntory	J	18	New York State University	US	6
Sumitomo	J	16	Agency of Industry, Science and Technology	J	5
Cetus	US	16	Salk Institute	US	5
Eli Lilly	US	15	Pasteur Institute	F	5
Shionogi	J	15			
Ashai	J	15			
Green Cross	J	15			

Includes all human DNA sequence patents granted worldwide. Where patent ownership is shared between two or more organizations, it is assumed that collaborators contributed equally. For the purpose of quantitative analysis, each party was therefore assigned an equal fraction of a patent. Countries of origin: J, Japan; US, United States; F, France; S, Switzerland.

institutions not been more active in applying for patents on human DNA? Many scientists in the public sector would argue that it is not their role to take the first steps towards commercialization of basic research. Further, some see the identification, isolation and cloning of human DNA sequences as pre-competitive. The traditionally poor links between universities and industry in the United Kingdom and other parts of Europe have also no doubt played a part in the low level of European public-sector patenting. In Japan, university researchers are not only prohibited by law from participating in research based in industrial laboratories, but government financial regulations inhibit interaction between universities and industry⁷.

So what human DNA sequences have

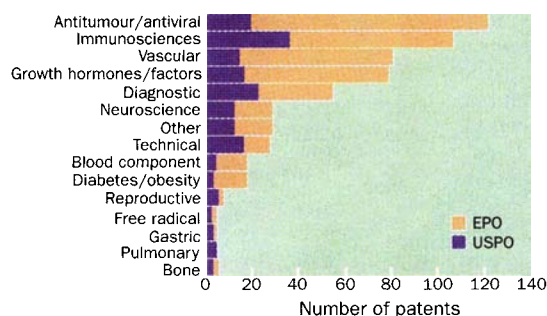


FIG. 2 Classification of human gene patents.

been patented? They range from primers for diagnostic use to patents for synthetic hybrids of interleukin and interferon genes. We divided the DNA sequences from the USPO and the EPO into 12 main categories (Fig. 2). The research areas covered by USPO patents are broadly similar in emphasis to those specified by EPO patents, although there is some variation. The category with the most patents overall is antitumour/antiviral, followed closely by immunosciences, growth hormones/growth factors and vascular categories. The combined first ranking of the antitumour/antiviral category emphasizes the importance of this research area for pharmaceutical companies; the category includes 58 EPO patents for sequences coding for interferons, interferon receptors and enzymes in the signalling pathway. And the dominance of the immunosciences in human DNA patenting is due to the early cloning successes with many natural and engineered genes for interleukins and their receptors. These products have therapeutic potential in stimulating the immune system and treating common allergies, as well as chronic diseases such as arthritis. About half the sequences in the EPO component of this category are for interleukin genes. Overall, the data show that the same therapeutic goals are being pursued by a wide range of companies.

Patenting DNA from any species is

currently acceptable under patent law provided that the usual criteria for novelty, non-obviousness and utility are met. Nucleic acids do not have any kind of exemption simply because they are the chemical compounds that form the basis of inheritance. As a result, more than a thousand patents for human DNA sequences have already been granted to nearly 300 organizations. The incentive to patent has been and continues to be industrial competitiveness for novel therapies and diagnostics based on specific DNA sequences. Industry, long aware of the critical importance of patenting in drug development, has provided the main impetus. Nevertheless, institutions in the public sector, particularly in the United States, are following the same path. The value of this approach

has been borne out by the judgement in *Amgen v. Genetics Institute*, which concluded that inclusion of DNA sequence information in patent claims was essential to secure intellectual property rights to erythropoietin, currently biotechnology's highest-earning drug at \$1.5 billion a year.

While Europe has over the past eight years debated the pros and cons of a harmonized directive on biotechnology intellectual property rights, Japan and the United States have between them secured rights to 70 per cent of EPO patents for human DNA sequences. That Europe's share of these EPO patents is so small is a major cause for concern. Foreign applicants are probably attracted to Europe because of its relatively new and less cumbersome EPO system, which processes patents within two years and is cheaper than the US system. Despite their vocal public opposition to biotechnology patents, European 'greens' have had little influence on patenting of DNA sequences (as opposed to plants and animals) in Europe.

What is not in question is European investment and competence in genomics. Although the European Union has spent some ECU75 million on genome projects, strong commitment to early release of sequence data into the public domain, rather than patenting, has been its priority. The shifting divide between public- and private-sector sequencing is nowhere more clearly illustrated than in the recent events surrounding the release of the full (that is, genomic) DNA sequence containing the breast cancer susceptibility gene *BRCA2*. Viewed as 'pre-competitive' information, and put on the Internet by the Sanger Centre at Cambridge, England, and the University of Washington, St Louis, it was rapidly exploited by the US company Myriad Genetics to locate the gene and file a patent for all diagnostic and therapeutic applications⁴. Putting sequence information in the

public domain at the earliest opportunity will promote research in the public sector and accelerate gene identification and, inevitably, subsequent patent filing by industry. Novel sequence information alone is pre-competitive and unpatentable. Combined with utility and inventiveness, it can qualify as a patentable invention. That a US genome company (a sector almost absent from Europe) is enabled to file for a patent by the agency of publicly funded research in the United Kingdom sits uncomfortably with UK government policy on wealth creation through basic science.

Challenges to utility have been largely confined to the controversy over the rejected patent application filed by the US National Institutes of Health in 1991 for 1,250 partial human gene sequences or 'expression sequence tags'. But the criteria for scope and inventiveness of patents have become hotly contested issues⁸. Given, as we see here, that DNA sequences are widely accepted as inventions and that patenting of human genes has been endorsed by the Human Genome Organization, it is unrealistic and impractical to seek alternative systems of protection such as copyright. Instead, attention and energy must be given to restricting the scope of patents so that the prospect of one company having proprietary rights over an entire gene and its mutations for all diagnostic and therapeutic purposes is no longer a reality. Public interest would be better served by policies encouraging competitiveness for improved second-generation healthcare products. Finally, it has been argued that it is obvious and no longer inventive to clone genes. In general, the law has been generous and ruled that cloning is indeed inventive. With the isolation of many new disease-related genes by the use of techniques of differential gene expression, the number of sequences for which patents will be filed will increase dramatically over the next three years. But as gene-cloning procedures become more standardized, patenting genes may well become more difficult. Rightly so. Meanwhile, Europe would do well to ponder over the stake that Japan and the United States already have in owning the human genome. □

S. M. Thomas, A. R. W. Davies, N. J. Birtwistle, S. M. Crowther (Science Policy Research Unit) and J. F. Burke (Department of Biochemistry) are at the University of Sussex, Falmer, Brighton BN1 9RF, UK.

1. Cole-Turner, R. *Science* **270**, 52 (1995).
2. Dixon, D. *Nature* **378**, 425 (1995).
3. *Science* **266**, 208-210 (1994).
4. Dixon, D. *Nature* **379**, 574 (1996).
5. Office of Technology Assessment *Biotechnology in a Global Economy* (OTA, Washington DC, 1992).
6. European Commission *The European Report on Science and Technology Indicators, 1994* (EUR15897, Brussels, 1994).
7. Swinbanks, D. *Nature* **371**, 731 (1994).
8. Thomas, S. M., Kimura, K. & Burke, J. F. *Res. Pol.* **24**, 645-663 (1995).

Raising dust in the greenhouse

Meinrat O. Andreae

The amount of dust in the atmosphere is increasing, largely because of human activity. Its climatic importance is becoming clear, as is our ignorance about aerosols in general.

JUST 25 years ago, it appeared we were rushing headlong into an ice age. During the previous 30 years, the mean temperature in the Northern Hemisphere had dropped by about 0.3 °C, and concern was raised that human activities might be to blame. In a 1971 paper in *Science*, Rasool and Schneider warned: "An increase by only a factor of 4 in global aerosol background concentration may be sufficient to reduce the surface temperature by as much as 3.5 °C. [Atmospheric aerosols consist of particles that are small enough to be suspended in the troposphere for days to weeks.] If sustained over a period of several years, such a temperature decrease over the whole globe is believed to be sufficient to trigger an ice age"¹. Concern about global cooling was soon abandoned, however, as worldwide temperatures rose sharply beginning in the 1970s, and the climatic effects of aerosols were all but forgotten by the early 1990s when the warmest years on record were observed. Attempts by a small group of scientists to remind their colleagues, and the world at large, of the climatic role of aerosols^{2,3} were dramatically underscored by the 1991 eruption of Mt Pinatubo, which released enough dust and sulphate aerosols into the atmosphere to reduce the global mean temperature by about 0.5 °C.

Two papers in this issue address an aerosol source that has been all but ignored in discussions of anthropogenic climate change: the flux of mineral aerosols from soils disturbed by erosion, over-cultivation and deforestation, mostly in arid or subarid regions. On page 419, Tegen *et al.*⁴ analyse the climatic impact of an estimated twofold increase in the global flux of mineral dust caused by human activity. Based on model calculations, they find that such an increase has pronounced but complex effects on climate forcing (a change of energy input to the climate). The paper by Li *et al.*⁵ on page 416 provides evidence from field measurements that dust aerosols indeed

play an important role in the atmospheric radiation budget.

The fundamental constraint on the Earth's energy budget is that the flux of incoming energy from the Sun, most of which is in the visible part of the spectrum, must be balanced by an equal outgoing flux in the form of infrared radiation. Any deviation from this balance drives the Earth's climate to a warmer or

light-scattering effect is well known to the air traveller who often finds it difficult to see the ground through a glaring layer of haze. Over industrialized parts of the world, sulphate particles produced by the oxidation of anthropogenic sulphur dioxide make up most of this scattering aerosol. In the tropics, biomass burning in forests and savannahs is the dominant source of aerosol particles, which consist mainly of organic matter and black carbon.

© 1996 EUMETSAT

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

A dust plume stretching from Africa to the Caribbean, photographed by Meteosat on 8 April 1994. Even in Barbados, at the far end of the plume, this dust reflects much more sunlight than other aerosols. The global climate is probably affected far more by dust from disturbed soils than has been assumed.

cooler state, with the rate of change determined by the magnitude of the imbalance and the inertia of the planet's heat reservoirs, particularly the oceans. 'Greenhouse' gases trap some of the outgoing infrared radiation, and thereby force the Earth's surface to come to a higher temperature so that the requirement for energy balance will again be fulfilled.

In contrast to the greenhouse gases, which act only on the outgoing, infrared radiation, aerosol particles can influence both sides of the energy equation, depending on their composition and size. Particles in the size range 0.1 to 1 µm are the most effective at scattering incoming solar radiation, and return some of it to space, thereby reducing the amount of energy available to warm the Earth. This

Until now, mineral dust has not been considered in most discussions of anthropogenic climate change, since it was considered to be part of the "natural aerosol background"⁶. Tegen and colleagues point out, however, that about half of the current dust loading is due to the disruption of soils by changing use of land in arid and subarid regions, such as the Sahel in Africa. This agrees with measurements of the amount of African dust transported over the Atlantic, which has increased dramatically in recent decades⁷.

The climatic effect of mineral dust particles is complex: because of their size distribution and chemical composition they scatter and absorb both incoming (visible) and outgoing (infrared) radiation. In the visible part of the spectrum, the light-scattering effect dominates, and the mineral dust aerosol has a net cooling effect, similar to that of sulphate and smoke aerosols. In the infrared region, however, mineral dust is an effective absorber and acts much like a greenhouse gas, trapping outgoing radiation and thus warming the Earth.

In their model calculations, Tegen *et al.* estimate that the backscattering of sunlight by dust and the additional trapping of outgoing infrared radiation by the same particles cancel almost exactly. That does not mean, however, that an increased dust loading has no effect on climate, because their conclusions indicate that the scattering of visible light by dust results in a reduction in the energy flux arriving at ground level of about 1 W m⁻², whereas the trapping of infrared radiation heats

the atmosphere. To put these effects in context, it is estimated that anthropogenic greenhouse gases at present exert a warming forcing of about 2.7 W m^{-2} , and that sulphate aerosols and biomass smoke probably have a cooling effect⁸ in the range 0.5 to 2.0 W m^{-2} . One may be tempted to conclude that, within the large uncertainties involved, cooling and warming by aerosols and greenhouse gases almost cancel each other at present. This simplification is, however, highly misleading.

We must remember that the influence of aerosols is highly variable in space and time. The sulphate aerosols are most abundant in the regions downwind of their sources, mostly in the industrialized areas of the Northern Hemisphere. Biomass smoke is present predominantly during the dry season in tropical areas. At a given aerosol concentration, its radiative effects are sensitive functions of solar angle, humidity and the albedo of the underlying surface, as well as the composition and size of the particles.

The regional variability of greenhouse and aerosol forcing is demonstrated impressively in Tegen and colleagues' Fig. 2 on page 421, which shows that strong cooling is concentrated over the Arabian Sea, the subtropical North Atlantic and the western North Pacific. The findings of Li *et al.* bear this out for the trade-wind region of the North Atlantic. They have measured the composition and optical properties of the atmospheric aerosol at Barbados, downwind and across the Atlantic from the dust sources in sub-Saharan Africa. They observe that, although dust is a much less effective scatterer than sulphate (per unit mass), its much greater abundance more than makes up for its weaker interaction with visible light. At Barbados, scattering by dust is four times that by sulphate. Even so, Tegen and colleagues' Fig. 2 shows that Barbados is actually at the far end of the African dust plume, and that large areas of the globe are subject to much stronger dust forcing still.

Correction

M. F. Perutz points out that, contrary to the statement in his News and Views article of 21 March (*Nature* **380**, 205–206; 1996), there is convincing spectroscopic evidence that dithionite does transfer NO from the thiol 93β to the iron in haemoglobin (J. Bonaventura, personal communication). Moreover a transient rise in diastolic blood pressure by 10 mm after transfusion with haemoglobin may be "not serious" for some patients, but might put others at risk. That risk could be avoided by the transfusion of S-nitrosohaemoglobin. On the first line of page 206, 'cysteine' should read 'cystine'.

The spatial structure of aerosol forcing leads to a highly uneven distribution of warming and cooling effects. And, since human beings and the biota experience not the global mean temperature *per se* but the effects of local climate and weather, this temporal and spatial variability is of central concern for the well-being of our planet's inhabitants. Tegen *et al.* have limited themselves in the present paper to calculating the forcing that results from adding dust to a climate model. They do not show the actual effects of this forcing on climate, nor have they allowed the model to react by changing atmospheric dynamics. Yet this is where more surprises may be in store. The shift of heating from the ground to the atmosphere, which results from the scattering of sunlight away from the ground combined with infrared absorption aloft, must change atmospheric stability on a large scale. That would affect atmospheric circulation patterns, which in turn would change the regional water cycle and lead to further changes in dust sources and sinks.

A full assessment of aerosol effects on climate requires their incorporation into fully interactive climate models, as Tegen *et al.* point out. This can lead to unexpected results. When the effect of sulphate aerosols on cloud optical properties was included in such a model, it predicted a pronounced winter cooling over the northern North Atlantic and a warming over central Europe, not at all what one would have anticipated from the source distributions⁹. Such a cooling over the North Atlantic had been an unexplained feature of the observational record of climate change. The distributions of predicted and observed climate change show significant correlations only if both greenhouse gas and sulphate aerosol effects are included in the models^{10,11}.

Far from reducing our concern over man-made climate change, our growing awareness of the climate effects of aerosols heightens it, as the inclusion of aerosol effects brings model predictions into better agreement with observations of climate change.

What consequences does this have for our expectations of future climate change? Paradoxically, the presence today of aerosol cooling increases the anticipated global warming in the next century. First, it had been argued that climate sensitivity (the amount of climate change for a given radiative forcing) was much lower than the models had predicted. This argument was based on the fact that the observed temperature change from the 19th century to the present had been at the very low end of model predictions. Inclusion of the aerosol effect has eliminated this discrepancy, supporting climate sensitivities around 3°C per doubling of atmospheric CO_2 (ref. 11). Second, because of the large inertia of the climate system, the effect of

greenhouse gas forcing takes decades to be fully transformed into climate warming, while the offsetting effect of aerosols acts almost immediately. Finally, the widely different atmospheric lifetimes of aerosols (days to weeks) and greenhouse gases (decades to centuries) will make their atmospheric concentrations take divergent paths in future decades. The short-lived aerosols are quickly removed and their concentration remains proportional to the concurrent rate of emission, while the greenhouse gases continue to accumulate in the atmosphere. As a result of these effects, a simultaneous sharp reduction of both CO_2 and aerosol emissions would actually lead to an increased rate of warming for some time¹².

I would like to conclude on an aspect most disturbing to an observational scientist. Today, we have outstanding documentation of the growth of greenhouse gases in the atmosphere over the last two centuries. For atmospheric aerosols, there is nothing of the kind. The record of fossil fuel burning implies, of course, that emissions must have increased, and local aerosol measurements over a few decades also show significant increasing trends, as in the case of mineral dust at Barbados⁷. Ice-core data also support at least regional increases in aerosol fluxes. But we are not able to prove, on the basis of direct observation, that the global atmospheric aerosol burden has increased since the beginning of the industrial revolution. We are not able to validate the model predictions of anthropogenic aerosol burdens by appropriate measurements. Even worse, we have no system of sufficient coverage and accuracy in place to allow scientists thirty years from now to assess whether aerosol burdens have changed since the present. Given the importance and complexity of aerosol effects on climate, it appears essential to put such an observational capability into place. □

Meinrat O. Andreae is in the Department of Biogeochemistry, Max Planck Institute of Chemistry, PO Box 3060, D-55020 Mainz, Germany.

1. Rasool, S. I. & Schneider, S. H. *Science* **173**, 138–141 (1971).
2. Charlson, R. J. *et al.* *Science* **255**, 423–430 (1992).
3. Charlson, R. J., Lovelock, J. E., Andreae, M. O. & Warren, S. G. *Nature* **326**, 655–661 (1987).
4. Tegen, I., Lacis, A. A. & Fung, I. *Nature* **380**, 419–422 (1996).
5. Li, X., Maring, H., Savoie, D., Voss, K. & Prospero, J. M. *Nature* **380**, 416–419 (1996).
6. Andreae, M. O. in *Future Climates of the World* (ed. Henderson-Sellers, A.) 341–392 (Elsevier, Amsterdam, 1995).
7. Prospero, J. M. & Nees, R. T. *Nature* **320**, 735–738 (1986).
8. Houghton, J. T. *et al.* (eds) *Climate Change 1994: Radiative Forcing of Climate Change* (Cambridge Univ. Press, 1995).
9. Erickson, D. J., Oglesby, R. J. & Marshall, S. *Geophys. Res. Lett.* **22**, 2017–2020 (1995).
10. Santer, B. D. *et al.* *Clim. Dynam.* **12**, 77–100 (1995).
11. Mitchell, J. F. B., Johns, T. C., Gregory, J. M. & Tett, S. F. B. *Nature* **376**, 501–504 (1995).
12. Wigley, T. M. L. *Nature* **349**, 503–506 (1991).

The long and the short of it

Rob Benne

THE study described by Polson *et al.* on page 454 of this issue¹ will interest two distinct constituencies: virologists and those working on RNA editing. It adds both to knowledge of hepatitis delta virus (HDV), a human pathogen, and of the enzyme double-stranded RNA adenosine deaminase (dsRAD). This enzyme is ubiquitous in multicellular organisms, and is implicated in editing RNA by converting adenosine to inosine, a rare nucleic-acid base. Polson *et al.* show that, at least *in vitro*, dsRAD is capable of editing the HDV antigenome. *In vivo*, editing leads to the production of a longer, functionally different form of the protein product, hepatitis delta antigen. This work follows other papers²⁻⁴ that have provided *in vitro* evidence for the involvement of dsRAD in adenosine-to-inosine RNA editing in an entirely different system — the production of subunits of glutamate receptor (GluR) channels in mammalian brain.

The adenosine (A) to inosine (I) editing saga began with the discovery that mRNAs for five of the 16 or more subunits that assemble into GluR channels in mammalian brain are edited at several sites⁵⁻⁷. Further research revealed that, before splicing, A's in pre-mRNA are converted into I's by oxidative deamination^{5,6}. This conversion is dependent on relatively short dsRNA structures in which the editing regions and complementary sequences in the downstream intron participate (see figure). Inosines are read as guanosines (G) by the translational machinery (see ref. 2), explaining the alteration of the corresponding amino-acid residues in the GluR subunits.

The changes affect the physiological properties of the receptor. Editing at the so-called Q/R site of the mRNA encoding GluR subunit B, at which a CAG (glutamine) codon is changed into CIG (arginine), drastically decreases the Ca²⁺ permeability of receptors containing subunit B; editing at the R/G site, by contrast, results in faster recovery rates from channel desensitization⁸.

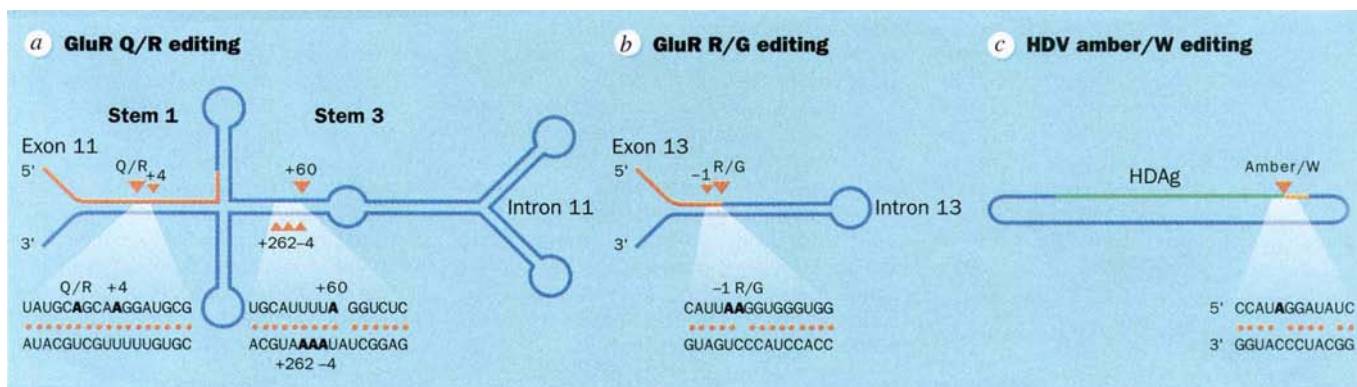
On the HDV side of things, study of RNA editing has up to now been the domain of virologists. HDV is a subviral human pathogen, the propagation of which depends on concurrent infection with hepatitis B virus⁹. The HDV protein product comes in two forms, depending on the editing at the so-called amber/W site, which converts a stop codon (UAG) into a tryptophan codon (UGG). The viral life cycle is complex, and it has been unclear whether the editing is the result of a uracil (U) to cytosine (C) transition in the genomic RNA strand or an A modification (to G or I) in the antigenomic strand. The confusion has been resolved in favour of the latter^{1,9}. The result of editing is a 19-amino-acid extension at the carboxy-terminus of the protein (hepatitis delta antigen, HDAg), which is again of biological relevance. The shorter HDAg protein is required for replication of the HDV genome, whereas the longer form inhibits replication and is required for packaging. HDV RNA editing, like that of GluR mRNA, requires base-pairing of the editing region (see figure).

It was the requirement for base-paired structures that made dsRAD, an enzyme found in the nuclei of many multicellular species⁶, a likely candidate for A to I edit-

ing. It was first discovered in *Xenopus* from its ability to unwind long dsRNAs by deaminating multiple A's to I's, resulting in the creation of unstable I:U basepairs¹⁰. *In vitro*, purified dsRAD deaminates multiple, randomly selected A's in long dsRNAs of more than 100 base pairs. The enzyme has little activity with short dsRNAs of 36 base pairs but, surprisingly, certain A's appear to be specifically deaminated (see ref. 1). The principal biological function of dsRAD is a matter for speculation, but it has been implicated in the modification of RNAs from defective measles viruses, implying that it is involved in cellular antiviral defence mechanisms.

Does dsRAD also have a role in editing? The studies¹⁻⁴ demonstrate that dsRAD is indeed capable of specific editing of A's in the short and irregular double-stranded structures in GluR-B mRNA and HDV antigenomic RNA, at least *in vitro*, but that the outcome differs between editing sites. The amber/W site of antigenomic (but not of genomic) HDV RNA¹ and the R/G site of GluR-B mRNA³ can be edited efficiently by dsRAD *in vitro*, resulting in editing patterns and frequencies similar to those found *in vivo*. Polson *et al.*¹ even find that purified enzyme by itself suffices for HDV RNA editing, no cellular or viral cofactors being required.

Editing of the GluR-B mRNA Q/R site presents a more complicated picture²⁻⁴. After incubating synthetic GluR-B RNA with purified mammalian recombinant dsRAD² or *Xenopus* dsRAD⁴, virtually no editing of the Q/R site (which is almost completely edited in rodent brain mRNA) was observed. In contrast, an intronic 'hotspot' 60 nucleotides downstream of the site was efficiently edited. The extent of editing at the Q/R site could be increased to up to 30 per cent by adding nuclear extracts of human embryonic



Adenosine-to-inosine editing sites in the messenger RNA encoding the B subunit of glutamate receptors (a, b), and the antigenomic RNA of hepatitis delta virus (HDV; c). a, Q/R editing. The pre-mRNA is edited at the Q/R and other sites (+4, +60 and +262-4) in exon 11 and intron 11. b, R/G editing at this and the -1 site in exon 13. Editing frequency at the different sites varies according to experimental protocols and the system studied (other sites than those mentioned here are edited at a low frequency). Edited A's are indicated by red triangles and by bold in the corresponding nucleotide sequences; exons

and introns are indicated by red and blue lines, respectively; dots indicate Watson-Crick or G:U basepairing. c, The HDV genome is a single-stranded circular RNA molecule which encodes a shorter or longer version of HDV antigen (HDAg) through editing of an antigenomic RNA. The figure shows the rod-shaped structure of the viral antigenomic RNA, and depicts the basepairing essential for editing at the amber/W site (red triangle). The reading frame encoding the shorter form of HDAg is depicted as a green bar; the extension originating from editing is orange.

kidney (HEK) cells, if dsRAD and extract were preincubated². Editing at the +60 site was drastically lowered by the addition of the extract, and Dabiri *et al.*² speculate that (protein) cofactors, present even in non-neuronal HEK cells that have very little intrinsic editing activity, modulate the accuracy of dsRAD-mediated editing. A similarly low level of Q/R site editing was found by Melcher *et al.*³, and the minor differences between the various sets of results probably stem from subtle differences in experimental design.

In all, it seems that, even in the presence of cofactors, *Xenopus* or mammalian dsRAD cannot catalyse efficient editing of the GluR-B mRNA Q/R site. In search of a dsRNA A deaminase that could, Melcher *et al.* screened a rat cDNA library with a DNA probe for the deaminase domain of dsRAD. They came up with a cDNA encoding a protein with a domain structure similar to dsRAD, sharing 31 per cent sequence identity as a result of a conserved carboxy-terminal domain. They showed that the encoded protein (RED1, for dsRNA-specific editase1) converts A's to I's in long duplex RNAs, is expressed in brain but also in other tissues and, most importantly, was capable of specific editing of the Q/R site *in vitro*. Superimposition of the *in vitro* dsRAD and RED1-specific editing patterns at and around the Q/R site resulted in a pattern closely resembling that in GluR-B mRNA from central neurons.

What do these results tell us? First, they allow another glimpse at possible structural requirements for specific A to I editing (but that is another story). Second, it is now clear that dsRAD is not the only dsRNA A deaminase out there — it could well be that the ubiquitous expression of dsRAD is in fact the result of a team effort by several family members. Given that most (pre)-mRNAs possess a substantial amount of secondary structure¹¹, this might imply that GluR and HDV RNA editing represent the tip of the iceberg and that many other mammalian RNAs are edited by A to I conversion. If this is so, it would certainly complicate the lives of those involved in human genome and expressed cDNA sequence projects.

A final issue is that of what the *in vitro* results mean *in vivo*. Could it be that RNA

editing is, in fact, the major biological function of dsRADs? Formal proof would require abrogation of enzyme activity by gene targeting or microinjected antibodies, a daunting task given the possible existence of several dsRADs with shared sequence elements and overlapping specificities and functions. But now that the first steps have been taken, and genome

and cDNA-sequencing projects can provide megabases of sequences at high speed, it has at least become feasible to look for other dsRNA A deaminases. □

Rob Benne is in the Department of Biochemistry, Academic Medical Centre, University of Amsterdam, Meibergdreef 15, 1105 AZ Amsterdam, The Netherlands.

QUANTUM PHYSICS

Discrete charm of the photon

Peter Knight

EVER since the work of Planck and Einstein at the beginning of this century, we have known that light fields come in discrete packages, which were only named 'photons' in the 1920s. Despite this being an accepted part of textbook quantum mechanics, experimental evidence of discrete quanta is hard to come by. Two papers just published back to back in *Physical Review Letters*^{1,2} now demonstrate this 'graininess' of quanta and open up new methods of testing the predictions of basic quantum mechanics — such as the persistence of quantum coherence for mesoscopic systems (the famous problem of Schrödinger's cat).

Radiation of frequency ω trapped in a cavity of volume V has an oscillating electric field that comes in discrete units with an amplitude of $E_0 = (\hbar\omega/2\epsilon_0 V)^{1/2}$, where ϵ_0 is the permittivity of free space. This quantum lump of electric field is not a negligible quantity: it is about a millivolt per centimetre in the cavities used in these experiments. If there are n quanta present, the quantum of electric field is $E_0 n^{1/2}$. But any quantum field is likely to be a superposition of states, and so to have a distribution of photon numbers. A perfectly coherent field has a Poisson distribution, whereas a thermal field has a distribution of the sort proposed by Bose and Einstein in the 1920s.

So how would you see the discreteness of the field? One way is to allow an atom to interact resonantly with the trapped radiation. The electric field does work on the atomic dipole so that the cavity field and atom exchange energy in a sinusoidal time-dependent evolution called a Rabi oscillation. At least that is what they would do if there were a precise number of quanta present. If the field contains a superposition of numbers n , then the evolution is a sum of discrete sinusoidal oscillations weighted by the probabilities of the different precise-number states. These will go out of phase so the coherence will be lost in the energy exchange — the famous 'collapse' characteristic of the quantum Rabi problem. But as the distribution is discrete the oscillations will partially revive (only partially,

because the characteristic frequencies are incommensurate).

Collapses and revivals are hard to see. Ordinary atomic states have electric dipole moments that are too small to see such coherent effects in a normal dissipative environment, where spontaneous emission smears out the very effects we want to see. But developments in quantum optics have allowed us to produce very highly excited 'Rydberg' atoms with giant dipole moments, which are nevertheless very stable, capable of resolving such effects provided the radiation itself is carefully contained in a cavity at low enough temperatures to shield the quanta from dissipation into the resistive cavity walls.

The pioneering experiment in this area was performed by Herbert Walther's group at the Max Planck Institute for Quantum Optics in Munich using Rydberg atoms to both create and probe radiation trapped in a superconducting cavity (discussed in ref. 3). The main source of field was the stimulating current of atoms injected into the cavity — this is the 'micromaser' or 'one-atom maser'. Each atom exits the cavity leaving the field somewhat modified from its previous state; the next atom sees this newly modified field. The collapse and revival of the transition probability seen is a consequence of the discreteness of the micromaser radiation field in the cavity.

The group of Serge Haroche in Paris has continued work on quantum features of cavity fields¹, again using Rydberg atoms to probe the field, while driving the cavity with an extremely weak, coherent photon field from an external source instead of pumping it with a beam of micromaser atoms. They are able to see precisely the effects of the discreteness of the radiation field in the collapse-revival dynamics. They have presented the measured spectrum of the quantum Rabi model showing exactly the discrete Rabi couplings expected.

At the same time the group of David Wineland in Boulder has shown how to study quantum effects of this kind without using a cavity². They use a single,

- Polson, A. G., Bass, B. L. & Casey, J. L. *Nature* **380**, 454–456 (1996).
- Dabiri, G. A., Lai, F., Drakas, R. A. & Nishikura, K. *EMBO J.* **15**, 34–45 (1996).
- Melcher, T. *et al.* *Nature* **379**, 460–464 (1996).
- Hurst, S. R., Hough, R. F., Aruscavage, P. J. & Bass, B. L. *RNA* **1**, 1051–1060 (1995).
- Scott, J. *Cell* **81**, 833–836 (1995).
- Bass, B. L. *Curr. Biol.* **5**, 598–600 (1995).
- Seeburg, P. H. *Trends Neurosci.* **16**, 359–365 (1993).
- Lomeli, H. *et al.* *Science* **266**, 1709–1713 (1994).
- Casey, J. L. & Gerin, J. L. *J. Virol.* **69**, 7593–7600 (1995).
- Bass, B. L. & Weintraub, H. *Cell* **55**, 1089–1098 (1988).
- Steitz, T. A. in *The RNA World* (eds Gesteland, R. & Atkins, J.) 219–237 (Cold Spring Harbor Laboratory Press, New York, 1993).

laser-cooled, trapped ion. The discreteness here is in the quantum vibrational motion of the ion — the coupling of the electronic excitation of the ion to its vibrational motion is of exactly the same mathematical form as in the atom-cavity field used by the Paris group, provided the ion sits at a node of the laser fields that excite its electronic transitions. But the dissipative effects are much less severe, and the ion motion is easier to manipulate than a cavity field.

The group first cool their beryllium ion to its lowest vibrational quantum state in the electromagnetic trap, and then proceed by a sequence of pulsed excitations of the driving laser to construct appropriate states. States with precise vibrational quantum numbers are first excited, and evolve sinusoidally as the electronic and the vibrational states exchange energy. As a by-product of this technique the group are able to see the motion decohere as the initial vibrational quantum number is increased. Highly excited quantum states of motion are more fragile and lose coherence far more rapidly to the environment, a phenomenon whose mechanisms are hotly debated by theorists.

To make a coherent excitation with a Poisson distribution of quantum number, analogous to the coherent field used in Paris, all they need to do is to displace the centre of the trap to effectively shake the ion into a classical oscillation. Again, once a coherent state is established the quantum Rabi evolution leads to a collapse and revival of coherence. The vibrational motion can even be squeezed — by parametric coupling (like a child amplifying a swing's motion by altering the centre of gravity, standing and sitting at twice the swing frequency), the distribution of vibrational quanta can be made narrower or broader.

Indeed, by careful choice of laser frequencies they are now able to engineer the way the ion sees its environment: different vibrational states can be excited so that the ion loses electronic excitations one at a time, or even two at a time. Environmental engineering will allow us to control the nature of the quantum state in equilibrium with the outer world. This opens up new avenues for the investigation of such hotly disputed areas as the decoherence problem, as well as for constructing highly entangled states of the kind needed to realize Schrödinger's cats and other animals so long a preserve of theory. □

Peter Knight is in the Department of Physics, Blackett Laboratory, Imperial College, London SW7 2BZ, UK.

Seeing cones in living eyes

Abbie Hughes

TEN years ago, a nineteenth-century mindset was broken by the prediction that human cones should be visible through the intact optics of the eye¹. Last month, the prediction's practical realization was reported by Miller *et al.*² who describe how they see cone photoreceptors, 3 μm in diameter, in the fovea of the eye. This remarkable technical feat foreshadows a powerful clinical tool.

Although it was technically possible for some 150 years, photoreceptors were seen through natural optics only recently.

self to the cone array, Helmholtz⁵ came to the same conclusion as Young: "the inexactness of the image... very little exceeds, under ordinary conditions of illumination, the limits which are set to the delicacy of sensation by the dimensions of the retinal cones". Both of these founders of physiological optics "take the optical image quality as a 'given' and assume that receptor grain adjusts itself accordingly"⁶.

The twentieth-century application of the Shannon-Nyquist sampling theorem to the retina consolidated their view. Consider the image of a black-and-white grating used in eye tests. A photoreceptor array can resolve the grating only when there is at least one independent receptor to detect each black or white bar. The theorem shows^{4,7} that if the optics can image a grating, or detail, finer than the receptor array, then it cannot be reported accurately and what is perceived will be irreversibly corrupted by undersampling. This is known as 'aliasing'. The nineteenth-century mindset thus assumed the 'grain' of image and cone array to be necessarily 1:1 'matched' for economy, an assumption endorsed in this century by the perceived need to avoid aliasing. Such matching would make it impossible to view the elements of the cone array with an ophthalmoscope. Researchers thus failed to look for cones or to interpret anything they did see as photoreceptors.

But is the assumption of 1:1 matching valid? Apparently not in insects⁷: their optical resolution is finer than the receptor arrays. Could this also occur in the vertebrate simple eye?

The answer is yes. The question of design principles for ideal vertebrate eyes was explored by Snyder *et al.*¹. Simple physical arguments and information theory show information transfer to be maximized when the grain of the optical image is double that of the receptor array. The penalty from consequent aliasing was insignificant. Desirable optical image quality is then^{1,6} "determined by a trade-off between the advantage of increased sensitivity on one hand,

IMAGE
UNAMAGABLE
FOR A CORRECT
FOR A CORRECT
REASONS

Thomas Young (1773–1829). Young's analysis of ocular imagery may have inadvertently set the stage for the misconception that photoreceptors are not visible through the optics of the living eye. Nonetheless, the breadth of insight in his 1801 paper *On the Mechanism of the Eye* has rarely been equalled. He is commemorated in Westminster Abbey as "Eminent in every department of human learning": the first to have "established the undulatory theory of light" and the first to have "penetrated the obscurity which had veiled for ages the hieroglyphics of Egypt".

Why? If they were observable, why had they not been serendipitously recognized? Perhaps Thomas Young set the stage in his 1801 classic, *On the Mechanism of the Eye*^{3,4}. He saw the gradient of refractive index in the lens as ensuring the best possible image; optics is given. The 'grain' of this image is then passively matched by an economical, 'just good enough', array of discrete light-sensitive elements in the retina.

It subsequently came to be realized that the light-sensitive elements are of two kinds, cone photoreceptors for daytime vision and rod photoreceptors for night vision. Although he confined him-

1. Brune, M. *et al.* *Phys. Rev. Lett.* **76**, 1800–1803 (1996).

2. Meekhof, D. M., Monroe, C., King, B. E., Itano, W. M. & Wineland, D. J. *Phys. Rev. Lett.* **76**, 1796–1799 (1996).

3. Knight, P. *Nature* **326**, 329 (1987).

weighed against the disadvantage of corruption due to aliasing on the other hand". The sampling theorem had set an artificial limit. The mindset was broken: optical 'grain' should generally be finer than the cone array.

So, when can receptors be seen? It isn't generally recognized that cones resolving a grating pattern through the optics, and resolving the same cones from outside the eye, are very different tasks. Each bar of the finest resolvable grating must be matched by one cone. But for the same cones and the 'gaps' between them to be resolved through natural optics, the image 'grain' must be at least twice as fine as the grain of the cone array⁶.

For this reason, analysis predicts that human cones should be resolvable through a 2-mm natural pupil¹. But in garter snakes the optical image quality is three times better than the resolving power of its cone array, and therefore ideal for testing the theory of receptor visibility¹. Cones were thus first seen in the intact snake eye⁸, followed by receptors in the toad eye⁹. But still, according to Miller *et al.*², "Despite such theoretical predictions, there is no compelling evidence that receptors can be resolved in the image of the living human retina". Their results now confirm theory; human cones are visible in the fovea.

To see them, the retina was illuminated by diffused, 555-nm laser pulses, of 4 ms duration to minimize blur from eye movements. The subject's pupil was dilated, accommodation paralysed and the retina imaged onto a CCD camera through a 5-mm artificial pupil, larger than in a conventional fundus camera to transmit more fine detail. Computer modelling of anatomical photoreceptor arrays imaged through the subject's optics generated patterns convincingly like those from the camera².

We are left with several questions. Why, for instance, is viewing technically easy in snakes and difficult in humans, where dilated pupils, correction of higher-order optical aberrations and computer enhancement may be needed? Are cones potentially visible through the optics of the eye anywhere in the retina? Can they be seen in all species?

If Young were correct there would be little variation in cone visibility across eyes or species. His unavoidable aberrations would similarly degrade the off-axis images of all vertebrate species. The

matched grain of the peripheral cone array would be equally coarsened. Such classic 1:1 matching in the periphery would prevent cones being seen anywhere in vertebrate eyes.

But Young was wrong. Comparative studies reveal many patterns of image quality and cone topography — visual streaks, even high-resolution areas far from the optic axis^{4,10}. Oblique aberrations are not inevitable because optics is a flexible element 'moulded' to local receptor grain^{1,4}. In contrast to nineteenth-century belief, receptor grain and its pattern is the given, not optics.

Realization of this flexibility opens up a new dimension in the analysis of eye design. The concept of receptor array 'grain' may be generalized for performance characteristics other than information transfer¹ — colour, motion-sensitivity⁷ or threshold detection at array frequency. Localized retinal specialization means that optical quality, cone array, the requisite matching and thus receptor visibility, may all vary across the visual field in accord with the underlying needs of the species in question. Other technical issues may require consideration, especially for viewing rods², but at least part of the cone array should be readily seen in the retina of any species¹.

Clinicians⁴ will soon be able to observe directly, at the receptor level and in living human eyes, both the progress of foveal disease and the course of its treatment. With current equipment not everyone is a good subject and the accessible retina is restricted. But adaptive optics should improve detail by minimizing higher-order aberrations and image enhancement will help detection of faint features². Confocal scanning laser ophthalmoscopy provides another route to receptor structure¹¹. Although it was invented for examining eye defects, the conventional ophthalmoscope turned out to be invaluable for diagnosis of systemic disease. Similar unpredictable benefits may result from routine high-resolution ophthalmoscopy.

Grappling with the theory of eye design has led to great practical advances. Theories can be surprisingly powerful influences. They determine mindset — what we look for and what we find. □

Abbie Hughes is in the Institute of Advanced Studies, Canberra, ACT 2601, Australia.

DAEDALUS

Common sense

In a democracy, the law is ultimately an expression of popular opinion. Every lawyer is supposed to compare his judgements with those of an imagined average citizen — 'the man on the Clapham omnibus' or 'John Q. Public'. A jury is intended to be a sample of such citizens; indeed a televised trial invites the whole audience to play at being jurors. The idea is to bring the common sense of the people to bear on the case.

In practice, of course, the law and common sense parted company long ago. Most businesses are enmeshed in a mass of daft legal regulations whose point, if they ever had one, has long been made irrelevant by changed circumstances. Modern absurdly over-protective product liability laws have wiped out whole industries. Most citizens see the law as an expensive nonsense to be avoided at all costs. Daedalus now wants to put the common sense back into it.

His weapon is DEX, the Democratic Expert System he outlined last week. This computerized database will accumulate public opinion on a huge range of issues, and will smooth and generalize it into an opinion map in 'issue space' which can be sampled at any point. Accordingly, it can predict the spread of answers to questions never asked by its pollsters. DEX is thus a unique repository of the 'common sense' of all the citizens. When it has been set up and is running well, Daedalus will use it to make judgements on real and hypothetical legal cases.

Lawyers will be outraged. They will never allow the judgements of DEX to have legal force, however well they represent popular opinion. But this may not matter. Suppose, for example, that some crazed bureaucrat decides to forbid the sale of cheese, on the grounds that it contains bacteria and thus contravenes the food-safety regulations. Suppose an armed robber sets out to sue the police for the emotional distress of his arrest. DEX might predict (say) that 95 per cent of the population would throw out these cases as idiotic. Even with no legal force behind them, DEX's rulings would have a powerful influence on any jury. They would make these legal actions too risky to pursue. Common sense would be back!

DEX will swiftly drain the legal swamp. Absurd regulations will lose their force without being repealed; nobody will even try to enforce them in the face of almost certain defeat in the courts. Better still, since DEX will give an opinion on purely hypothetical cases, potential litigants could check out their chances without going near a lawyer. Even the few cases worth pursuing will be cheaply settled out of court, on whatever basis DEX regards as fair.

David Jones

1. Snyder, A. W., Bossomaier, T. J. & Hughes, A. *Science* **231**, 499–501 (1986).
2. Miller, D. T., Williams, D. R., Morris, G. M. & Liang, J. *Vision Res.* **36**, 1067–1079 (1996).
3. Young, T. *Phil. Trans. R. Soc. Lond.* **B92**, 23–88 (1801).
4. Hughes, A. in *Visual Neuroscience* (eds Pettigrew, J. D., Sanderson, K. J. & Levick, W. R.) 60–89 (Cambridge Univ. Press, 1986).
5. Helmholtz, H. v. *Popular Scientific Lectures* (ed. Klein, M.) 116–117 (Dover, New York, 1962/1881).
6. Snyder, A. W., Bossomaier, T. J. & Hughes, A. in *Vision: Coding and Efficiency* (ed. Blakemore, C.) 45–52 (Cambridge Univ. Press, 1990).

7. Snyder, A. W. in *The Handbook of Sensory Physiology*, Vol VII/6A *The Physics of Vision in Compound Eyes* (ed. Autrum, H.) 225–314 (Springer, Berlin, 1979).
8. Land, M. & Snyder, A. W. *Vision Res.* **25**, 1519–1523 (1985).
9. Jagger, W. S. *Vision Res.* **25**, 729–731 (1985).
10. Hughes, A. in *The Handbook of Sensory Physiology*, Vol VII/5 *The Visual System in Vertebrates* (ed. Crescitielli, F.) 615–756 (Springer, Berlin, 1977).
11. Campbell, M. C. W., Munger, R. & Stargard, M. *Optom. Vis. Sci.* **70** (12s), 52 (1993).

Origin of the arthropod mandible

SIR — Arthropods, vast in number and with enormous variation in body forms, are a fascinating group. We have found that myriapods (millipedes, centipedes and allies) have different mandibular origins from insects and crustaceans, which is of consequence for resolving phylogenetic relationships among major groups of arthropods.

For more than a century, the phylogenetic relationships among the main arthropod lineages have been a topic of lively discussion. Almost every imaginable combination has been proposed, but at present only two hypotheses are seriously considered: the 'TCC' view, which separates trilobites, crustaceans and chelicerates from the rest of the arthropods¹, and the 'mandibulate' theory, which groups together crustaceans, insects and myriapods². One feature is common to both: the close relationship between myriapods and insects. These two arthropod groups were traditionally united into Atelocerata³, because they share five adult characteristics: a tracheal system, malpighian tubules, absence of appendages corresponding to the second antennae of crustaceans, unbranched legs, and a mandible (jaw) composed of a whole limb.

The existence of the Atelocerata has recently been questioned by two independent studies reporting molecular data to infer arthropod phylogeny^{4,5}. Friedrich and Tautz⁴ suggested that crustaceans, and not myriapods, are the sister group of insects, arguing that the first three characteristics common to both myriapods and insects are convergent adaptations to terrestrial life and thus do not reflect a common ancestry. Panganiban *et al.*⁵ have shown also that differences between branched and unbranched legs can be caused by a simple developmental switch. In the light of these arguments, it is essential to evaluate critically the fifth feature shared by myriapods and insects, namely, a similarity in the structure of their mandibles.

Traditionally, crustacean mandibles are regarded as being formed from a limb base, and ateloceratan mandibles as being composed of a whole limb⁷, but this view has been challenged recently². One way to resolve this dispute is to investigate the structure of arthropod mandibles at the molecular level, as suggested by Friedrich and Tautz⁴.

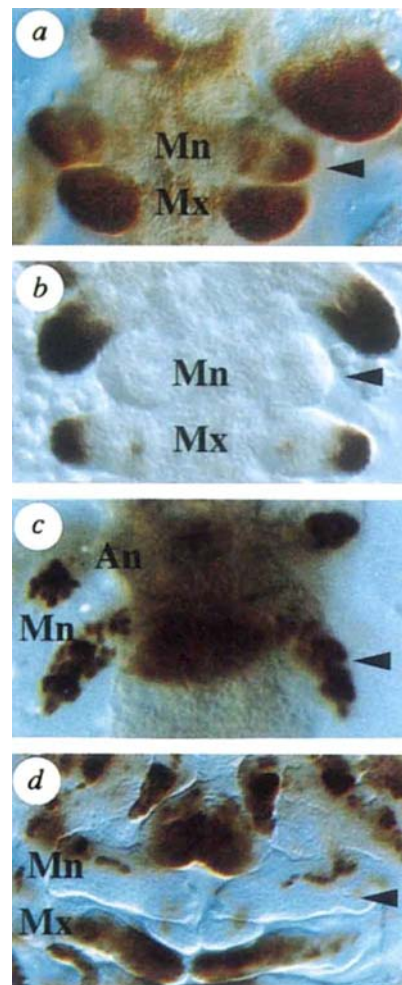
We studied the pattern of expression of the homeobox gene *Distal-less* (*Dll*) in the mandibles of the millipede, and compared it with results from insects and crustaceans (*Dll* antibody was kindly provided by G. Panganiban). This gene specifies the distal part of appendages, and therefore can be used as a molecular marker

for investigating the structure of the mandibles (whole limb versus limb base only). In the millipede *Oxidus gracilis* (a in the figure), *Dll* is expressed in the distal part of the mandibles, as predicted in ref. 4, indicating their whole-limb structure.

In light of the recent discovery of the Cambrian fossil whose head and trunk appendages were long and leg-like⁸, the whole-limb mandibles of today's myriapods probably represent an ancestral arthropod state. Thus, we have a testable hypothesis: if myriapods and insects are indeed sister taxa, then *Dll* should also be expressed in insect mandibles. But Panganiban *et al.*⁹ have shown that *Dll* is not expressed in mandibles of modern insects. To examine whether the absence of *Dll* is characteristic of the whole insect lineage, we included in our analysis the primitively wingless insect *Thermobia domestica*, and found that *Dll* is not expressed in the mandibles of this species (b in the figure); this further suggests that insect mandibles are formed from the limb base and may be similar to the mandibles of adult crustaceans.

Both *O. gracilis* and *T. domestica* undergo direct development, where immature stages differ from the adults mainly in the development of the gonads and genitalia. This allows us to correlate directly embryonic changes in *Dll* expression with structural changes in adult mandibles. In contrast, only crustaceans that undergo larval development have been studied so far⁶. Consequently, the finding that *Dll* is expressed throughout the mandibles of crustacean nauplius larvae (c in the figure) is not informative because the larval cells expressing *Dll* do not contribute to the adult structures (as noted in ref. 6). To infer the origins of the mandibles in adult crustaceans, it is necessary to study species that undergo direct development. We therefore included the terrestrial isopod *Armadillidium vulgare*, a direct developer, in our analysis. We found that crustacean mandibles are indeed composed of a limb base only, as is evident by the lack of *Dll* expression in ectoderm (d in the figure).

In summary, our data are consistent with earlier predictions^{2,4,6} that the arthropod mandible was originally composed of a whole limb and was similar to the present-day mandibles of the myriapods. Further, our data suggest that during arthropod evolution, the mandible structure changed from a whole limb (a in the figure) to a limb base only, the latter type being a shared feature between insects and crustaceans (b, d). This finding has two important implications. First, it argues against the traditional view that insect and crustacean mandibles are fundamentally different⁷; and second,



Expression pattern of *Dll* in the mandibular segments of the millipede *Oxidus gracilis* (a), the primitively wingless insect *Thermobia domestica* (b), the crustacean nauplius larva *Artemia franciscana* (c) and the terrestrial isopod *Armadillidium vulgare* (d). An, antennal segment; Mn, mandibular segment; Mx, maxillary segment. Arrowheads indicate the mandibular appendages.

it directly supports the hypothesis that crustaceans, not myriapods, are the sister group of insects⁴. The overall similarity of nervous and visual systems in both insects and crustaceans provides independent support for this hypothesis¹⁰. These findings lessen the case for uniting insects and myriapods into Atelocerata.

**Aleksandar Popadić, Douglas Rusch
Michael Peterson, Bryan T. Rogers
Thomas C. Kaufman**

Howard Hughes Medical Institute,
Department of Biology,
Indiana University,
Bloomington, Indiana 47405, USA

1. Cisne, J. L. *Science* **186**, 13–18 (1974).
2. Kukalová-Peck, J. *Can. J. Zool.* **70**, 236–255 (1992).
3. Telford, M. J. & Thomas, R. H. *Nature* **376**, 123–124 (1995).
4. Friedrich, M. & Tautz, D. *Nature* **376**, 165–167 (1995).
5. Boore, J. L. *et al.* *Nature* **376**, 163–165 (1995).
6. Panganiban, G. *et al.* *Science* **270**, 1363–1366 (1995).
7. Manton, S. M. *The Arthropoda: Habits, Functional Morphology and Evolution* (Clarendon, Oxford, 1977).
8. Robison, R. A. *Nature* **343**, 163–164 (1990).
9. Panganiban, G. *et al.* *Curr. Biol.* **4**, 671–675 (1994).
10. Osorio, D. *et al.* *Trends Ecol. Evol.* **10**, 449–454 (1995).

Bacterial infection and coral bleaching

SIR — Bleaching in stony corals is the result of disruption of symbioses between the coral hosts and photosynthetic microalgal endosymbionts (zooxanthellae). Coral bleaching of unprecedented frequency and global extent were reported in the 1980s and early 1990s¹⁻³. This process may be evoked by various environmental stimuli, including increased seawater temperature⁴ and ultraviolet radiation⁵. There has been speculation that large-scale bleaching episodes are linked to global warming^{6,7}. The data presented here demonstrate for the first time that coral bleaching, in this case, bleaching of *Oculina patagonica*, is caused by a bacterial infection.

We first observed bleaching of *O. patagonica* in the summer of 1993. Bleached patches of polyps appeared throughout the affected colonies (*a* in the figure). Although the tentacular rim of the polyp may retain its pigmentation, the extratentacular coenosarc of the tissue was bleached, showing a total loss of pigmentation (*b* in the figure). Histological sections of bleached tissues showed a 70–90% reduction of algal densities. The initial observation that led us to consider that coral bleaching may be due to bacterial infection was the presence of large aggregates of rod-shaped bacteria on the border of the bleached zone at the tentacular rim (*c* in the figure). In unbleached tissues, no bacterial aggregates were observed.

We collected bleached and unbleached corals from the Mediterranean coast of

Israel and 'streaked' samples onto marine agar. Characteristic cream-coloured bacterial colonies, which we call strain AK-1, appeared on all eight samples from bleached corals and were absent from all 14 unbleached corals. Strain AK-1 has been classified as a *Vibrio* by standard microbiological tests⁸.

Inoculation of healthy *O. patagonica* with pure cultures of *Vibrio* AK-1 resulted in bleaching. We performed two types of experiment. First, 5×10^6 bacteria were placed on each of five healthy corals, and the corals then put back in an aquarium maintained at 25 °C. All the corals showed bleaching after 6–8 days. In a control experiment in which sterile medium was used in place of the bacteria, none of the corals showed any signs of bleaching. Bleaching was due to the bacterial cells themselves, and not extracellular products, because ultrafiltration led to an inactive cell-free supernatant fluid. The washed cells were as active as the original culture.

The remaining experiments were carried out at 26 °C without removing the corals from two-litre aerated aquaria. In the first test, we added 10^5 *Vibrio* AK-1 per ml to two aquaria containing *O. patagonica*. Antibiotics were added to one of these infected aquaria. A third control aquarium was treated in exactly the same manner, except that we added sterile medium in place of bacteria. After 44 days, 90% of the coral surface in the experimental aquarium was bleached, whereas no bleaching was observed in the

control and the antibiotic-treated aquarium (*d* in the figure). In a second test, three aquaria were infected with 5×10^6 *Vibrio* AK-1 per ml and two additional aquaria served as controls. In all three experimental aquaria, there was extensive bleaching after 42 days and tissue retraction after 52 days. The two control colonies remained healthy. Attempts to infect corals with 10^5 *Vibrio* AK-1 per ml at 16 °C were unsuccessful.

Our findings demonstrate that the causative agent of bleaching of *O. patagonica* is *Vibrio* AK-1, which was present in all bleached *O. patagonica* examined, obtained in pure culture and caused bleaching when inoculated onto healthy (unbleached) corals. Seawater temperature is a contributing factor; an increase in temperature may influence the outcome of the infection by lowering the resistance of the coral and increasing the virulence of the bacterium.

A. Kushmaro, Y. Loya

M. Fine

Department of Zoology,

E. Rosenberg

Department of Molecular Microbiology
and Biotechnology,

Tel Aviv University,

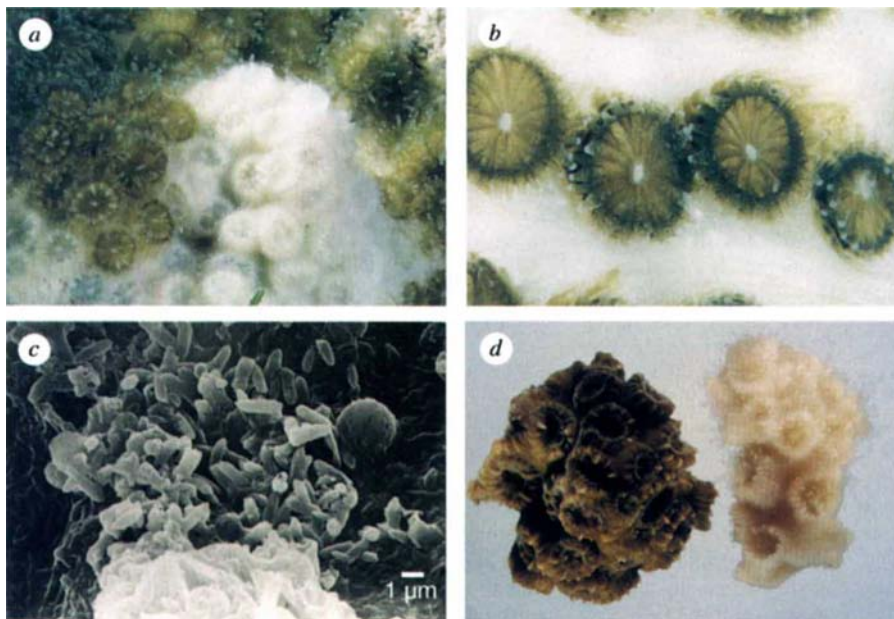
Ramat Aviv 69978, Israel

1. Williams, E., Goenaga, C. & Vicente, V. *Science* **238**, 877–878 (1987).
2. Brown, B. E. (ed.) in *Coral Reefs Spec. Iss.* **8**, 153–232 (1990).
3. Glynn, P. W. *Trends Ecol. Evol.* **6**, 175–179 (1991).
4. Jokiel, P. L. & Coles, S. L. *Coral Reefs* **8**, 155–162 (1990).
5. Gleason, D. F. & Wellington, G. M. *Nature* **365**, 836–838 (1993).
6. Smith, S. V. & Buddemeier, R. W. *A. Rev. Syst.* **23**, 89–118 (1992).
7. Glynn, P. W. *Coral Reefs* **12**, 1–17 (1993).
8. Farmer, J. J. III & Hickman-Brenner, F. W. in *The Prokaryotes* 2nd edn (eds Balows, A. et al.) 2952–3011 (Springer, New York, 1992).

Pest control by fluorescence

SIR — Environmentally safe biological control agents, such as baculoviruses, are among the most promising alternatives to synthetic chemical insecticides. Among the deterrents to their commercial development has been the lack of cost-effective procedures for monitoring their efficacy and ecology in nature. Green fluorescent protein (GFP) is a very stable protein which emits green light on excitation with ultraviolet or blue light at 395 or 470 nm (ref. 1). We have transferred the GFP gene from the jellyfish *Aequorea victoria* into a typical baculovirus, the *Autographa californica* multiple nuclear polyhedrosis virus (AcMNPV)², thus producing an easily visible marker for detecting infected insects.

We excised a complete GFP-coding sequence from plasmid pGFP (Clontech) and inserted it into the polyhedrin gene-containing transfer vector pAcUW21



The coral *Oculina patagonica*. *a*, Bleached patches of polyps throughout the affected colony ($\times 2$). *b*, Extratentacular coenosarc of the bleached tissue showing a total loss of pigmentation ($\times 5$). *c*, Scanning electron microscopy showing large aggregates of rod-shaped bacteria on the border between the bleached and unbleached zones at the tentacular rim. *d*, Control: healthy coral (left) and a bleached coral (right) 44 days after infection with bacterial strain AK-1 ($\times 2$). Photographs by A. Shoob.

(PharMingen) under the control of the strong P10 promoter of AcMNPV². We constructed the recombinant virus by co-transfection of the transfer vector and AcMNPV DNA into *Spodoptera frugiperda* (SF21) cells and easily isolated it by detecting the emission of green fluorescence from the infected cells.

AcMNPV has a wide host range among lepidopterous pests, including the diamondback moth *Plutella xylostella* L., and we used this species to illustrate the application of this novel GFP-detection system. This moth has in recent years become the most destructive insect of cruciferous plants and is the most universally distributed of all the Lepidoptera throughout the world. It has developed insecticide resistance in many countries, which has encouraged further interest in its biological control³.

The occlusion body, which contained infectious GFP gene-containing AcMNPV, was fed to the early third-instar larvae of the diamondback moth. Within 3–5 days of infection at 26 °C, we observed strong green fluorescence in the infected

larvae on irradiation by a hand-held long-wavelength (365 nm) ultraviolet light (Fig. 1). On average, fluorescence was detected 4 days before the death of infected larvae at 26 °C. Fluorescence patterns, either spotted or completely covering the organism, were observed, depending on the progression of viral infection (Fig. 2). The fluorescence lasts for at least 5 hours under continuous ultraviolet irradiation, or for more than 10 hours under intermittent exposure, without significant loss of fluorescence intensity. Addition of substrate is not necessary for light emission, and the strong fluorescence can be detected for at least 4 days at 22 °C or for 6 days at 4 °C in live larvae. At 4 °C the fluorescence can still be strongly excited 10 days after the first detection of fluorescence, even though the larvae may have died after 6 days.

The fluorescence is very intense and can be used for the early identification of infected individuals among large populations of healthy larvae (Fig. 2). Future developments should result in a useful, previously unavailable tool for precise and convenient assessment of infected larvae on crops. By using this tool, researchers can avoid collecting and recording only those insects showing severe symptoms of viral infection while missing those symptomless insects with disease still in progress. These severely infected larvae usually drop and dissolve within a relatively short time and are thus difficult to monitor on crops. Another advantage of the GFP recombinant virus is that there is no need for tedious molecular analyses to determine whether the death of insects was caused by the virus artificially released for insecticidal purposes or by natural pathogens.

To improve the efficacy of baculovirus as a pathogen for killing insect pests, the baculovirus genome has been modified in several laboratories by incorporating insect-specific toxins, hormones or enzymes^{4,5}. The similarly designed GFP gene-containing baculovirus can be used to predict the dispersal and/or long-term persistence of the recombinant viruses before their widespread use, which may adversely affect the environment. Because only a portable, long-wavelength

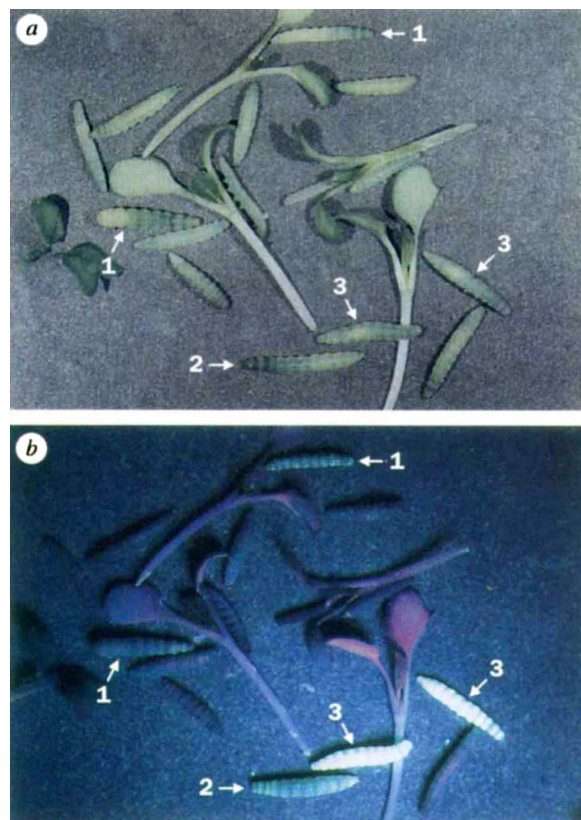


FIG. 2 Detection by ultraviolet light of diamondback moth larvae infected with the GFP gene-containing baculovirus. One group of larvae was infected with the recombinant baculovirus for 4 days at 26 °C, while the other group was not infected. The two groups of larvae were mixed together and photographs were taken using either visible (a) or 365-nm ultraviolet (b) light. Larvae at a relatively early stage of viral infection are labelled '1', those at an intermediate stage of infection are labelled '2' and those at a late stage of infection are labelled '3'. The uninfected insects are not labelled.

ultraviolet light source is required for demonstrating the strong, long-lasting emission of green fluorescence from the infected larvae, exciting displays can be made to increase the public's awareness of the benefits of genetic engineering.

Yu-Chan Chao

Su-Liang Chen*

Chih-Fen Li

*Institute of Molecular Biology,
Academia Sinica, Nankang,
Taipei, Taiwan, Republic of China*

*Also at Institute of Biotechnology, Chinese Cultural University, Taipei, Taiwan, Republic of China.

1. Prasher, D. C. *Trends Genet.* **11**, 320–329 (1995).
2. O'Reilly, D. R. *et al. Baculovirus Expression Vectors* (Freeman, Oxford, 1992).
3. Talekar, N. S. & Shelton, A. M. *A. Rev. Ent.* **38**, 275–301 (1993).
4. Tomalski, M. D. & Miller, L. K. *Nature* **352**, 82–85 (1991).
5. Cory, J. S. *et al. Nature* **370**, 138–140 (1994).

Nature regrets that in the Scientific Correspondence "Renal abnormalities in mutant mice" by C. Carpenter *et al.*, on page 292 of last week's issue, the figure was reproduced in black and white instead of colour.

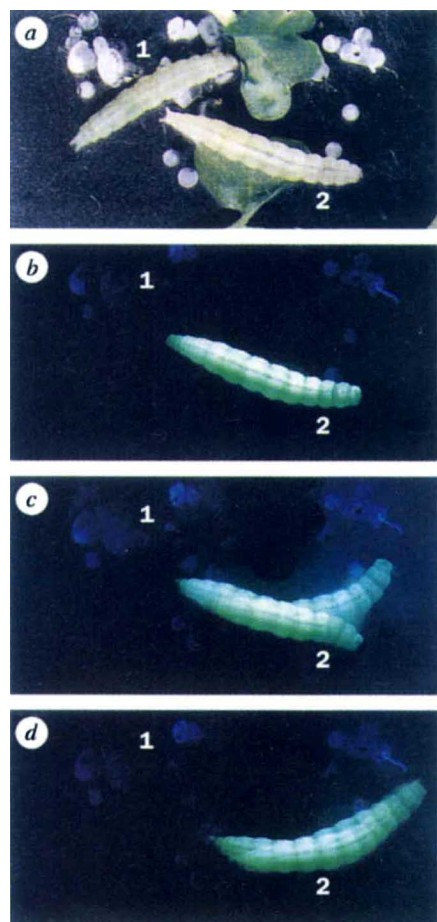


FIG. 1 Baculovirus-mediated expression of green fluorescence in diamondback moth larvae. a, Photograph of a non-infected (labelled '1') and an infected ('2') diamondback moth larvae taken in visible light. b–d, The same larvae photographed under long-wave ultraviolet light. These successive photographs show the movements of fluorescence-emitting larva ('2').

Human parasite finds taxonomic home

SIR—The phylogenetic analysis of ribosomal RNAs from *Blastocystis hominis* provides another interesting example of how molecular data can elucidate the taxonomic affinity of an organism that has proved intractable by traditional techniques. *B. hominis*, an obligately anaerobic protist inhabiting the human intestine, has been a taxonomic enigma since its initial description more than 100 years ago¹. It has variously been (mis)classified as the cyst form of a flagellate, an amoeba, a yeast and a sporozoan, due to the presence of multiple morphological forms in infected individuals¹⁻³. No flagellated stage has ever been identified in *Blastocystis*: it possesses mitochondria with tubular cristae, and its most characteristic feature is a crescent cap of heterochromatin in the nucleus. Because electron microscope studies have failed to place *Blastocystis* in any particular phylogenetic group, we undertook a molecular systematic study of *B. hominis* and a *Blastocystis* isolate from the guinea pig.

We sequenced the complete *Blastocystis* small-subunit ribosomal (r)RNA genes and aligned them with the homologous coding regions from organisms represent-

ing all main eukaryotic lineages. The two *Blastocystis* isolates are most closely related to each other; high bootstrap values in parsimony and distance analyses unequivocally place *Blastocystis* within the stramenopiles (see figure).

Stramenopiles, together with green plants, red algae, animals, fungi and alveolates, separated nearly simultaneously approximately one billion years ago⁴. Molecular phylogenies define stramenopiles as a complex and heterogeneous evolutionary assemblage that includes unicellular and multicellular protists, with both heterotrophic and photosynthetic representatives^{5,6}. Brown algae (kelp), diatoms, slime nets and water moulds are just a few examples of the diverse organisms found in this group. All stramenopiles possess mitochondria with tubular cristae and, unlike all other eukaryotes, have tripartite tubular hairs either on their cell surface or, more commonly, on their long anterior flagellum.

Blastocystis is a very unusual stramenopile. Its mitochondria do possess tubular cristae, but, because *Blastocystis* is anaerobic, the metabolic functions of the mitochondria are unclear (the mitochondria lack many elements of the classical

energy-producing pathway)¹. *Blastocystis* also lacks flagella and flagellar hairs. This probably results from a secondary loss of these features, as its closest relatives, the Proteromonadidae and Bicosoecida, have flagella and tubular hairs.

Despite continuing controversy over details of the *Blastocystis* life-cycle, this organism shares certain life-history traits with its sister taxon, *Proteromonas*. Both organisms are gut endosymbionts of vertebrates and encyst to an environmentally resistant form that appears to allow transmission between hosts^{7,8}. A prevalent *Blastocystis* morphotype has a membrane-bound 'central body' in which progeny appear to form by schizogony⁷, reminiscent of the multiplicative cysts of *Proteromonas*⁸.

Molecular analyses may provide insights into the recognition of new *Blastocystis* species and their host range. The small-subunit rRNA genes from the two *Blastocystis* isolates differ by 6.4% (see figure). This evolutionary distance is comparable to that observed between disparate animal⁹, plant⁹, fungi¹⁰ and stramenopile¹¹ species. Additionally, immunological and DNA hybridization analyses^{12,13} have revealed variation among *B. hominis* isolates from humans, raising the possibility that more than one species can infect a host. This new information about the phylogenetic affinity of *Blastocystis* provides the first demonstration of a stramenopile that infects humans.

Jeffrey D. Silberman

Mitchell L. Sogin

Center for Molecular Evolution,
Marine Biological Laboratory,
Woods Hole, Massachusetts 02543, USA

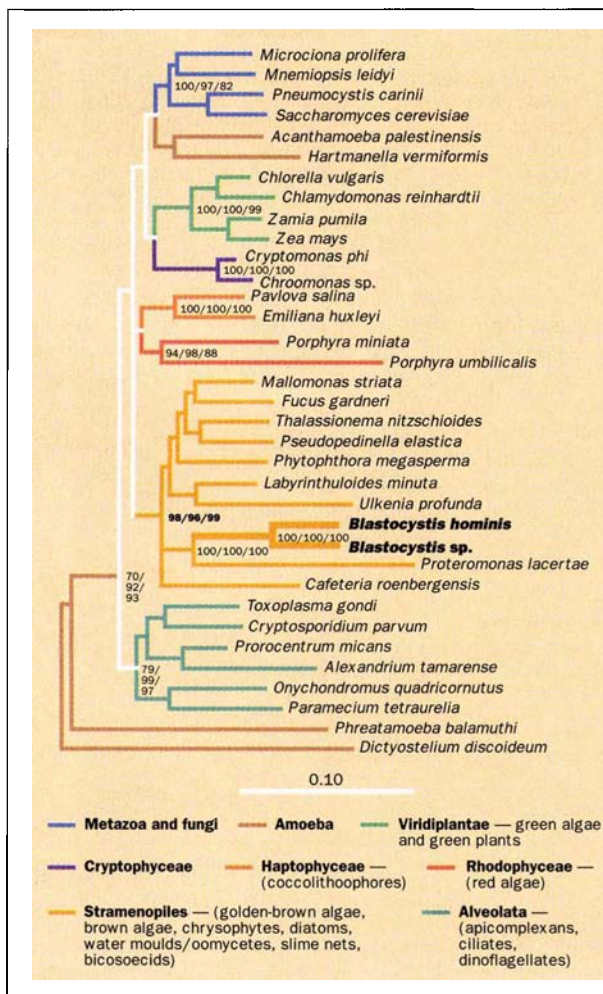
Detlef D. Leipe

NCBI (GenBank), National Library
of Medicine,
National Institutes of Health, Building 38A,
8600 Rockville Pike,
Bethesda, Maryland 20984, USA

C. Graham Clark

Department of Medical Parasitology,
London School of Hygiene
and Tropical Medicine,
Keppel St, London WC1E 7HT, UK

Relationship of *Blastocystis* species among eukaryotes as determined by phylogenetic analyses of 16S-like rRNA gene sequences. *Blastocystis* (ATCC 50177 and 50578) 16S-like rRNA genes were amplified *in vitro* from genomic DNA, cloned, sequenced and aligned with homologous rRNA coding regions from other eukaryotes. Phylogenetic reconstructions using maximum-likelihood, Fitch-Margoliash, neighbour-joining and parsimony analyses show that *Blastocystis* is embedded within the stramenopile lineage (only values from 100 bootstrap replications of the latter three analyses supporting the major eukaryotic groups are shown). *Blastocystis* sp. ATCC 50578 was cultured from guinea pig faeces provided by C. Chrisp, Unit for Laboratory Animal Research, Ann Arbor. Sequences have been deposited with GenBank; accession numbers are: *B. hominis*, U51151; and *Blastocystis* sp., U51152. The aligned dataset is available on request.



A biological Sonderweg

Paul Weindling

Biologists Under Hitler. By Ute Deichmann. *Harvard University Press*: 1996. Pp. 468. \$45, £28.50.

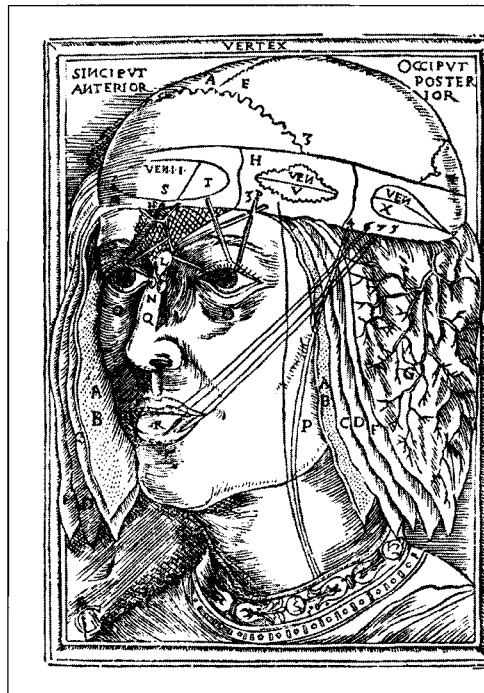
ALTHOUGH for Hitler politics was applied biology, the ability of Nazi leaders to regiment biologists was arguably more limited than often assumed. That the biochemist Otto Warburg (despite part-Jewish ancestry) and the Russian radiation geneticist Nikolai Timoféef-Ressovsky flourished in Hitler's Berlin indicates that there were at least pockets of innovative research. Given that the president of the Deutsche Forschungsgemeinschaft (DFG, literally 'German Research Foundation') from 1937 was an SS *Brigadeführer*, its biological funding raises intriguing problems. Ute Deichmann finds that not only did funding for biological research increase under Nazism, but also that "the influence of politics and ideology on biology in Germany was negligible". She denies that the Nazi demand for an action-oriented biology amounted to very much, or that the dismissal of Jewish and other critical biologists impeded research.

Much depends on the definition of biology. Deichmann classifies as full-fledged biologists those who gained the German higher doctorate (*Habilitation*) and had positions in universities, or who researched in an institute of the prestigious Kaiser Wilhelm Gesellschaft. Her classification excludes biologists in medical faculties or in other state institutes, field stations and museums (such as the dissident museum curator and hydrobiologist Walther Arndt, who was executed). One searches in vain for Albrecht Hase, an entomologist who switched from his university position of 'habilitated' zoologist to work on pest eradication with Zyklon B (prussic acid) at the Biologische Reichsanstalt, or for zoologists employed by medical faculties, such as Hermann Poll (who committed suicide after his dismissal in 1933). Deichmann discusses the lives of individual biochemists, but neither do they count as part of the cohort. There is, moreover, little justification for ignoring professional associations of botanists, zoologists and geneticists, such as the Deutsche Gesellschaft für Vererbungswissenschaft, which provide a historically more authentic representation of the population of biologists as a whole.

Max Delbrück condemned German biology for its "organisational rigidification". Encountering such creatures as "associate professor without civil servant status" suggests that here was a rigid status-bound system lacking in innovative

capacity. Perhaps the narrow definition of a field requiring commitment to generalized teaching of biology resulted in the marginalization of innovative researchers in biochemistry and genetics even before the Nazi persecutions. Deichmann does not take into account the earlier handicaps of Jewish scientists who had since the

Deichmann concludes that there was no correlation between funding of DFG applications and commitment to the National Socialist German Workers' Party and the SS. In doing so, she somehow forgets the notorious example of the link between the human geneticist Otmar von Verschuer and the notorious Nazi doctor Josef Mengele, a collaboration that was part of a DFG project. Although Hitler insisted on funding only those projects relevant to the war, Deichmann contends that "normal science" prevailed. But the racial priorities of war and conquest required an effective research community: whatever the personal politics of workers in virology, this subject area was a priority because of the



HEAD start — Johan Eichmann, a professor of medicine and mathematics at Marburg, used mediaeval sources for this illustration in his treatise *Anatomiae hoc est, corporis humani dissectionis pars prior* of 1537. Some of his other drawings were influenced by the perceptual art of the renaissance. Both can be found in the second edition of *An Illustrated History of Brain Function: Imaging the Brain from Antiquity to the Present* by Edwin Clarke and Kenneth Dewhurst. Norman, \$135.

1890s developed such fringe subjects as biochemistry.

Deichmann identifies only ten biologists (Friedrich Brieger, Gottfried Fraenkel, Fabius Gross, Mathilde Hertz, Hans Kalmus, Hugo Merton, Tibor Peterfi, Ernst Pringsheim and Johannes Holtfreter) who emigrated to or via the United Kingdom: her criteria exclude the successful geneticists Charlotte Auerbach or Hans Grüneberg. The United Kingdom avidly devoured not generalists but those with technical skills in areas such as genetics, bacteriology and biochemistry. I would include in the select group of 'biologists' who came to the United Kingdom a further 27 biochemists, and to these should be added some mercurial organic chemists. That recordings of birdsong by the musician Ludwig Koch struck a chord with British ornithologists exemplifies the creative synthesis that human migration could bring about.

epidemic hazards of the war in the East. Other DFG applications mentioned a commitment to research for the projected African colonies, and cancer research was seen by senior members of the Nazi party as of great importance in promoting the health of the racial élite. Biologists did not have to join the renegade Ernst Lehmann's nationalist Biologenverband to support the Nazi sanctioned movement for an applied, action-oriented biology; Deichmann's conceptualization of biology is excessively positivistic.

In-depth case studies reveal how scientists, who enjoyed considerable postwar renown, exploited research opportunities provided by the Holocaust. Konrad Lorenz was involved with the Posen Population Selection Office and the geneticist Hans Nachtsheim experimented on children from a euthanasia institution. Such biographies undermine the scientometrics. The SS sponsored as part of a geno-

cidal masterplan apparently innocent projects such as an expedition to Tibet and the cultivation of herbs and frost-resistant plants. The Kaiser Wilhelm Institute for Biology competed with the Kaiser Wilhelm Institute for Breeding Research and the SS for control of botanical institutes in occupied areas of the Soviet Union. Deichmann confirms an already established historical picture of SS science as ideologically diverse and innovative in such areas as genetics and epidemiology. Her claim that biologists were better off in Nazi Germany than in the Soviet Union overlooks the fact that a working research community was vital for Nazi racial policies. Research applications are in themselves a poor indicator of ideology and complicity in genocide — here one must reconstruct the broader thought-world of biologists and turn to science in action.

Although Deichmann portrays German biology as a scientific autarky, international contacts should not be underrated. Links continued throughout the war with France, and the unoccupied zone provided contacts to the United States which remained neutral until late 1941. (Deichmann mentions the US interest in 1941 in the work of Gerhard Schramm on nucleic acid and virus protein.) Links were sustained with Switzerland, Spain and Sweden throughout the war, aiding access to UK and US scientific periodicals. Deichmann's prime indicator for the postwar

period is the Science Citation Index. Given that only 39 German periodicals were sources for the period 1945–55 as opposed to more than 350 US periodicals, the more frequent citation of émigré biologists than those in Germany is hardly surprising.

Deichmann's revisionist verdict is that emigration did not account for the German failure to contribute to molecular biology, despite the loss of biochemists, virologists and protein chemists. Moreover, German wartime research achievements were as nothing compared to the Allied successes with penicillin, atomic energy, DDT and new vaccines. The German research effort was impeded by professors engaged in petty rivalries. Research may have thrived, but this was because of its complex links with Nazi priorities and authorities. Deichmann believes that the German failure to channel its research effort into molecular biology was due to the moral failure of biologists rather than any external political forces. The physiologist A. V. Hill reflected that scientists need a Hippocratic oath for dealing with social pressures. In Nazi Germany, this oath might have improved research quality more than the oath of loyalty to the Führer. □

Paul Weindling is at the Wellcome Unit for the History of Medicine, University of Oxford, 45 Banbury Road, Oxford OX2 6PE, UK.

but immunohistochemically appears not to be co-localized with it in cells. The two findings could be reconciled if the immunohistochemical results were due to masking of the reactive epitope of calponin in the contractile domains, perhaps by actin, rather than to absence of the protein.

The book's multiple personality is split between reference text, introductory monograph, methods handbook and a compilation of recent developments by specialists. Much in its pages will be of use to those either working or merely interested in the field. A review in *Nature* of an earlier opus on smooth muscle heralded it "From Hypertension to Erection"; is it a sign of progress that the present volume includes a cure for impotence? For those interested, it is to relax venous smooth muscle in the appropriate places by increasing its cyclic GMP content.

The review of protein phosphatases, including the recently identified trimeric myosin-bound phosphatase, and the chapter on energetics and several others manage to be as up to date as the printed word can be today. Some interesting developments since 1994 — evidence of the rapid growth in this field — could naturally not be included. For example, contractile regulation through caldesmon phosphorylation by MAP kinase seems less likely now, and the role of conventional and novel protein kinase Cs in G-protein-coupled inhibition of myosin light-chain phosphatase is being downgraded. We also learned that the 'duty cycle' of cross-bridges is longer in smooth than striated muscles, allowing the maintenance of higher force with less myosin.

No book review is respectable without a few quibbles. The erroneous dismissal (in two chapters) of the likely regulation of crossbridge cycling kinetics by Mg-ADP is not corrected by the editor, a distinguished biochemist of the classical school, who certainly knows that ADP and ATP occupy the same nucleotide-binding site on myosin and, therefore, that such effects of ADP will depend not only on its cellular concentration but also on the relative concentrations of the two competing nucleotides and their affinities for actomyosin. Many would also disagree with the implication that filamentous actin contains calcium rather than magnesium. (Journalistic integrity compels me to express a personal interest in both of these quibbles.) Finally, the phosphorylation of smooth-muscle myosin by Calmodulin kinase II is overlooked.

On balance, the book is a good read and recommended, although not for a single evening. □

Andrew P. Somlyo is in the Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, Virginia 22908, USA.

Relaxed, not impotent

Andrew P. Somlyo

Biochemistry of Smooth Muscle Contraction. Edited by Michael Barany. *Academic*: 1996. Pp. 418. \$115.

If no good deed goes unpunished, then having spared the readers by not contributing to this volume, my punishment to review it is not too harsh — reading it was even pleasant. The book delivers more than promised by the title, unless electrophysiology of channels, the subject of two chapters, is now within the purview of biochemists displaced by the march of molecular biology and X-ray crystallography. In truth, the inclusion of motility assays, molecular biology and the wealth of information gained from permeabilized smooth muscles illustrates the increasingly interdisciplinary approach required for understanding the relationships between molecular mechanisms and 'real life'. A multi-author volume such as this benefits from many different points of view (or prejudices) to which the reviewer adds his own.

Few important aspects of smooth muscle are left uncovered, and those neglected, such as ultrastructure, are ade-

quately reviewed elsewhere. The 30 chapters cover signal transduction, including electromechanical and pharmacomechanical coupling, regulatory and putative regulatory proteins, and the 'big three' of contraction: myosin, actin and tropomyosin. The 'how to' sections, albeit sporadic, include methods for isolating proteins and measuring their phosphorylation, without overwhelming one with exquisite details of patch clamping. The balanced discussions of myosin and its isoforms and of the actomyosin ATPase cycle in solution could have benefited from consideration, even if brief, of the biochemical consequences of mechanical strain on crossbridge cycling in muscle. Concerning the thin-filament-associated proteins, caldesmon and calponin, one tends to agree that "direct evidence for the role of caldesmon is at present rather unsatisfactory", rather than with the assertion, in another chapter, of its "vital role". The difficulty, not unique to this field (for it applies to signal transduction in general), of relating *in vitro* biochemistry to physiological function is illustrated in the case of calponin. This actin-binding protein inhibits actomyosin ATPase in solution,

Acknowledging our relatives

Mary Midgley

The Others: How Animals Made Us Human. By Paul Shepard. *Island/Shearwater*: 1996. Pp. 374. \$24.95.

PAUL Shepard's theme is a good one. It is that our distancing of our own species from all other animals deprives us of a vitally necessary range of symbols. Because we no longer take other animals seriously, we now have an impoverished and distorted notion of humanity. Our absurd claim to stand outside the biosphere, whether as detached spectators or as engineers controlling it, deprives us of our foothold on reality. As he says:

Neither the creationists nor the postmodern critics are right, one thinking that the world was made for us and the other that it is made by us. A better vision of the animals is that we are one among them. The only drama in town... is the evolutionary play on the ecological stage.

The sense of continuity with other creatures that most early humans undoubtedly felt played (says Shepard) an essential psychological role. It provided a language for expressing and illuminating the complications of human social life. He points out how deeply this repertoire still pervades our speech today:

We... carp, rat, crow or grouse vocally.
We cow, quail, toady, lionize and fawn....
We bull, ram or worm our way, monkey with things, weasel and chicken out. We know loan sharks, possum players and bullshitters.

Even today, our small children are passionately and spontaneously interested in animals, even when discouraged from this by their elders. As for early humans, their cosmologies and origin-myths draw no sharp line around the human species at all. Most of them claim animal ancestry without embarrassment.

Is all this — as Enlightenment wisdom tells us — just useless childish luggage that we ought to lose now that we have made our journey to civilization? No, says Shepard. Our alienation from continuity with the beasts flows partly from our deliberate insistence on control rather than on communication, and partly from our ignorance:

The true richness of these social/natural metaphors depends on knowledge of the natural behaviour of the animals themselves, on the diversity and complexity of those animals, a habit of watching we have lost by stages. We disavowed the animal mind and soul three centuries ago, reduced the beasts to representa-



PECKING order — A brown goshawk (*Accipiter fasciatus*) makes a meal of a large rabbit. Adult females have been known to take prey up to twice their own weight whereas males confine themselves to victims of a far more modest size. Just one of many striking images to be found in *Australian Birds of Prey* by Penny Olsen. Johns Hopkins University Press, \$49.95.

tions of the appetites twenty centuries ago, replaced the wild forms in our environment with simpler domestic animals of fewer species sixty centuries ago.

Like wild herbs in cooking and medicine, notions of tapir ancestors and parrot paradigms have the quaint air of rustic stories... Yet in a time of deeply felt loss of 'nature', which we mistakenly think of as a kind of pastoral anodyne, we may wonder about the necessity of 'thinking' animals. [Shepard then quotes Terence Turner] "The 'nature' incarnated in animal symbols is not simply... a convenient metaphor for human social patterns... It consists of aspects of human society that are rendered inaccessible to social consciousness as a result of their incompatibility with the dominant social framework. These alienated aspects of the human (social) being, which may include the most fundamental principles of social and personal existence, are therefore mediated by symbols of an ostensibly asocial or 'natural' character."

In short, if we think about people in animal terms we can grasp essential things about them that we do not know yet and could not envisage if we tried to reach them directly. These terms provide the language for new insights that cannot penetrate our existing models of human psychology. Those models are misleading partly because they are far too simple, but also because — still worse — they are biased by all sorts of past choices of which we are unaware, choices that do not really suit our present and future at all.

What does all this amount to? I think Shepard is right to say that irrational insistence on human uniqueness distorts our thought on many important subjects. We do deeply need a more serious attention to our continuity with other animals.

When he is in a psychologically perceptive mood, he has some really excellent, original things to say about this, things that make the book well worth reading. It is, however, fearfully uneven, diffuse, sometimes eccentric and shockingly unreliable about details. As a quite typical instance, I was surprised to see myself quoted in large print at the head of a chapter as having said about animals that "They too are persons". This was in fact a remark I had made about unrelated *human beings*, in a short section clearly and unmistakably headed "Why Do Other People Matter?"

This oddity forms part of a savage campaign that Shepard wages against proponents of animal rights. He is wedded to the idea that contemporary hunters understand animals far better than humanitarian people do — in fact, that the only true way to understand animals is actually to hunt and eat them. This has to be a muddle. No doubt primitive subsistence hunters do understand animals intimately, and also have interesting ambiguous attitudes to certain animals whom they both revere and eat. But how much have these hunters in common with urban people who drive out in cars to shoot creatures for sport? Shepard's indiscriminating polemic damages his book because it is so unconstructive. We cannot now become unspoiled hunters and gatherers. We have to develop attitudes towards animals that make sense from where we now stand. Humane feeling is certainly not the only kind of material we need for these attitudes. But it is surely an essential part of it. □

Mary Midgley is at 1A Collingwood Terrace, Newcastle upon Tyne NE2 2JP, UK.

Fertile addition

Roger Gosden

Dynamics of Human Reproduction: Biology, Biometry, Demography. By James W. Wood. *Aldine de Gruyter*: 1995. Pp. 653. DM159, \$79.95 (hbk); DM80, \$39.95 (pbk).

CONSIDERING our teeming numbers, it seems odd that we are so infertile as a species, but that is nonetheless true — comparatively speaking. A late puberty, lack of synchrony between coitus and ovulation, long inter-birth interval and a pregnancy wastage rate of around 50 per cent mean that successful pregnancies are few and far between. When and how often to reproduce are two of the most important decisions in life and, through evolutionary selection, we and our great-ape cousins have adopted an unhurried pace, for there is time and opportunity for repeated breeding attempts. Matters are rather different for ticks, termites and even turtles that live in more hazardous circumstances. They have to invest earlier and more heavily in breeding to ensure the survival of their lines. In Ogden Nash's words:

*The turtle lives 'twixt plated decks
Which practically conceal its sex.
I think it clever of the turtle
In such a fix to be so fertile.*

Behind these generalizations there is a rich tapestry of variation in human reproduction, and this forms the substance of James Wood's contribution. At heart is the concept of 'natural fertility', originally adumbrated by the French demographer Louis Henry in the 1950s and 1960s. The expression is easily misunderstood, and should not be equated with the biological potential to reproduce, or fecundity. Rather, it is fertility in the absence of deliberate control, and is modified by both physiological effects, such as breast-feeding, and cultural obligations, including sexual abstinence.

Gathering reliable population data for studying natural fertility is getting harder by the year because of the acculturation of traditional societies. Henry relied heavily on historical data from his own country to estimate natural fertility. Others have studied religious isolates such as the Amish and the Hutterite communities of North America or tribal societies such as the !Kung of Botswana and the Gainj of the highlands of Papua New Guinea (the author's speciality).

The plethora of variables potentially affecting fertility must be reduced to a manageable size to avoid becoming bogged down. Wood adopts the proximate determinants of fertility because they are computable variables, such as age at the

menarche, marriage and menopause on the one hand, and the frequency of insemination, amenorrhoea and pregnancy loss (what he calls susceptibility factors) on the other. This is a fruitful strategy for mathematical modelling (often associated with John Bongaarts); and Wood uses proximate determinants as his rubric throughout the book because they mediate many fascinating influences on fertility. For instance, seasonal variation is still a mysterious influence, and peak fertility has been stable over historical time in some places, although no global pattern exists. What is more, the widespread assumption that the effects of seasonality or under-nutrition are simple adaptive traits for optimizing fertility do not bear close inspection. Age-specific fertility rates in females are more clear-cut because, irrespective of their levels, they decline steadily between the twenties and late forties in all populations.

Models of human reproduction have thrown light on many patterns and processes in reproduction — the fertile time of the menstrual cycle, influences of breast-feeding and nutrition, fetal loss,

onset of permanent sterility and so on. But many gaps still need to be filled, especially in our knowledge of the male contribution and the role of genetic variation. Methodological difficulties abound and there are pitfalls for the unwary, which prompted the statistician Udny Yule to tease his colleagues nearly a century ago: the correlation that existed between the frequency of sightings of storks in English counties and the numbers of human births did not necessarily imply a direct causal influence.

A scholarly integration of the facts of reproductive biology and the figures of demography in some 650 pages packed with illustrations is an ambitious task for the most indefatigable of monograph writers, but Wood has done the job admirably well. His book will appeal mostly to graduate students and researchers in demography; and, apart from squirming at the occasional fact, I enjoyed it, as will anyone engaged in human reproductive biology. □

Roger Gosden is at the Research School of Medicine, University of Leeds, Leeds LS2 9NS, UK.

The fascination of infestation

George C. McGavin

Insects Through the Seasons. By Gilbert Waldbauer. *Harvard University Press*: 1996. Pp. 289. \$24.95, £15.95.

GILBERT Waldbauer is one of those few lucky people paid to pursue their hobby. Reading *Insects Through the Seasons*, one discovers why he finds entomology endlessly fascinating. The large and beautiful saturniid moth *Hyalophora cecropia* that enthralled and intrigued the eight-year-old Waldbauer became one of his main research interests and is the star of the book. Details of its life cycle are used as stopping-off points to explore the many varied aspects of insect biology. Sex, violence and a cast of billions: what more could one possibly want? And as if his words, a blend of science and sentiment, were not enough to bring the subject to life, a cecropia moth flies across the bottom corner of the book as one flicks the pages.

Here readers will discover strange stories and fantastic facts about the lives of insects and the many ways in which millions of years of evolution have equipped these organisms, arguably the most successful on our planet, for survival. We also learn that evolution has provided other animals, and humans too, with some similar solutions to the problems of life in a hostile world.

In places I was left wanting a little more detail or explanation. There can be no doubting the benefits of complete

metamorphosis as witnessed by most insect species that develop in this way, but why did it evolve? Is gradual metamorphosis really a compromise and did complete metamorphosis arise because it would be successful? Is there the taint of teleology here? But these are minor quibbles when viewed against the book as a whole.

When we try to bend the natural world to suit our own needs we can solve some problems, such as the control of the terrible screw-worm fly by the release of sterile males. But things can, of course, go wrong: for the sake of a little more honey we now have to cope with murderous killer bees.

Predation and parasitism, the author reminds us, are recurring themes in the insect world. And while most insects attack each other in an awesome variety of seemingly unpleasant ways, some turn their attention to humans. We are told of one medical entomologist who carefully recorded, in gruesome detail, the entire 50 days and 15.5 hours taken by two bot fly maggots to develop inside his own body. But the quest for knowledge does not normally demand such dedication. The book encourages us to take a close look at nature in general and insects in particular. □

George C. McGavin is at the Oxford University Museum, Parks Road, Oxford OX1 3PW, UK.

Characterization and inhibition of a cholecystokinin-inactivating serine peptidase

Christiane Rose*, Froylan Vargas*, Patricia Facchinetti[‡], Pierre Bourgeat*, Ramesh B. Bambal[†], Paul B. Bishop[†], Suzanne M. T. Chan[†], Andrew N. J. Moore[†], C. Robin Ganellin[†] & Jean-Charles Schwartz*

* Unité de Neurobiologie et Pharmacologie (U. 109) de l'INSERM, Centre Paul Broca, 2ter rue d'Alésia, 75014 Paris, France

[†] Department of Chemistry, University College London, Christopher Ingold Laboratories, 20 Gordon Street, London WC1H 0AJ, UK

[‡] Faculté de Pharmacie, Université R. Descartes, 4 Avenue de l'Observatoire, 75006 Paris, France

A cholecystokinin (CCK)-inactivating peptidase was purified and identified as a membrane-bound isoform of tripeptidyl peptidase II (EC 3.4.14.10), a cytosolic subtilisin-like peptidase of previously unknown functions. The peptidase was found in neurons responding to cholecystokinin, as well as in non-neuronal cells. Butabindide, a potent and specific inhibitor, was designed and shown to protect endogenous cholecystokinin from inactivation and to display pro-satiating effects mediated by the CCK_A receptor.

CHOLECYSTOKININS (CCKs) constitute a family of hormonal and neuronal peptides that exert pleiotropic biological effects in gut and brain. One example, CCK-33, the sulphated tritriacontapeptide, is implicated in the control of gall-bladder contraction, gastric emptying, and intestinal motility¹.

In cerebral neurons, CCK immunoreactivity corresponds mainly to the sulphated carboxy-terminal octapeptide (CCK-8)^{2,3}, which fulfills all the criteria of a neurotransmitter^{4,5}. Remarkably, CCK immunoreactivity and dopamine coexist in mesolimbic neurons, and may be implicated in psychotic disorders⁶.

The actions of CCK are mediated by CCK_A and CCK_B receptors⁷, and various antagonists have helped to improve our understanding of the physiological role of CCK, for example in the control of food intake, which is enhanced by CCK_A antagonists⁸, and the control of anxiety, which is decreased by CCK_B antagonists⁷. However, no CCK agonist with acceptable bio-availability has yet been designed, despite the interest of such agents as research tools and potentially as drugs with pro-satiating and antipsychotic activity.

Another approach to this problem would be to design selective peptidase inhibitors to protect the endogenous neuropeptide against inactivation and thereby amplify the biological responses it triggers. The feasibility of this has been demonstrated for endogenous enkephalins with the identification of the inactivating ectopeptidase 'enkephalinase' (neprilysin, EC 3.4.24.11)⁹, the rational design of inhibitors^{10,11} that protect enkephalins¹² and their use in gastroenterology¹³.

In the case of CCK, however, the difficulty has been the identification of physiologically relevant peptidase(s) from a variety of candidates shown to cleave CCK-8 *in vitro*^{14–21}. As a result of studying the inactivation of endogenous CCK-8 released from depolarized brain slices, we proposed that one or more serine peptidases, which cleave the two peptide bonds where the carboxyl group is donated by a methionine residue (Fig. 1a), are the physiologically relevant peptidases^{22–24}. This conclusion was based on the identification of CCK-5 (Gly-Trp-Met-Asp-Phe-NH₂; GWMDF-NH₂) and Gly-Trp-Met (GWM) as major CCK-8 fragments, and the observation that serine reagents greatly

enhance recovery of released CCK-8 in intact form, which, in their absence, is less than 20%.

CCK-hydrolysing activities

It was not clear from previous studies^{22,23} whether one or several distinct ectopeptidases were responsible for the two cleavages leading to formation of CCK-5 and GWM. We therefore purified the enzyme(s) responsible using three parallel assay systems, to follow cleavages of the Met 3–Gly 4 and/or Met 6–Asp 7 amide bonds.

GWM formation was evaluated using a radioimmune assay (RIA)²⁵ and either the full CCK-8 molecule or its N-terminal hexapeptide (DYMGMW) as substrates to assess both cleavages together, or the Met 3–Gly 4 cleavage alone, respectively. The Met 6–Asp 7 cleavage was also selectively evaluated using CCK-5 (GWMDF-NH₂) as the substrate and a high-performance liquid chromatography (HPLC) assay for GWM.

Membranes of rat cerebral cortex were treated with Brij 35 and the resulting solubilisate submitted to a four-step HPLC purification during which the GWM-forming activity was enriched up to 8,600-fold. At each step the three peptidase assays led to overlapping chromatographic profiles (Fig. 1b), indicating that a single enzyme is responsible for both cleavages of CCK-8. This enzyme corresponded to a serine peptidase of relative molecular mass 135,000 (*M_r* 135K) that was irreversibly labelled on sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE) by [³H]di-isopropylfluorophosphate (DFP); this labelling was prevented in the presence of an inhibitor known to block endogenous CCK-8 inactivation in brain slices²² (Fig. 1e).

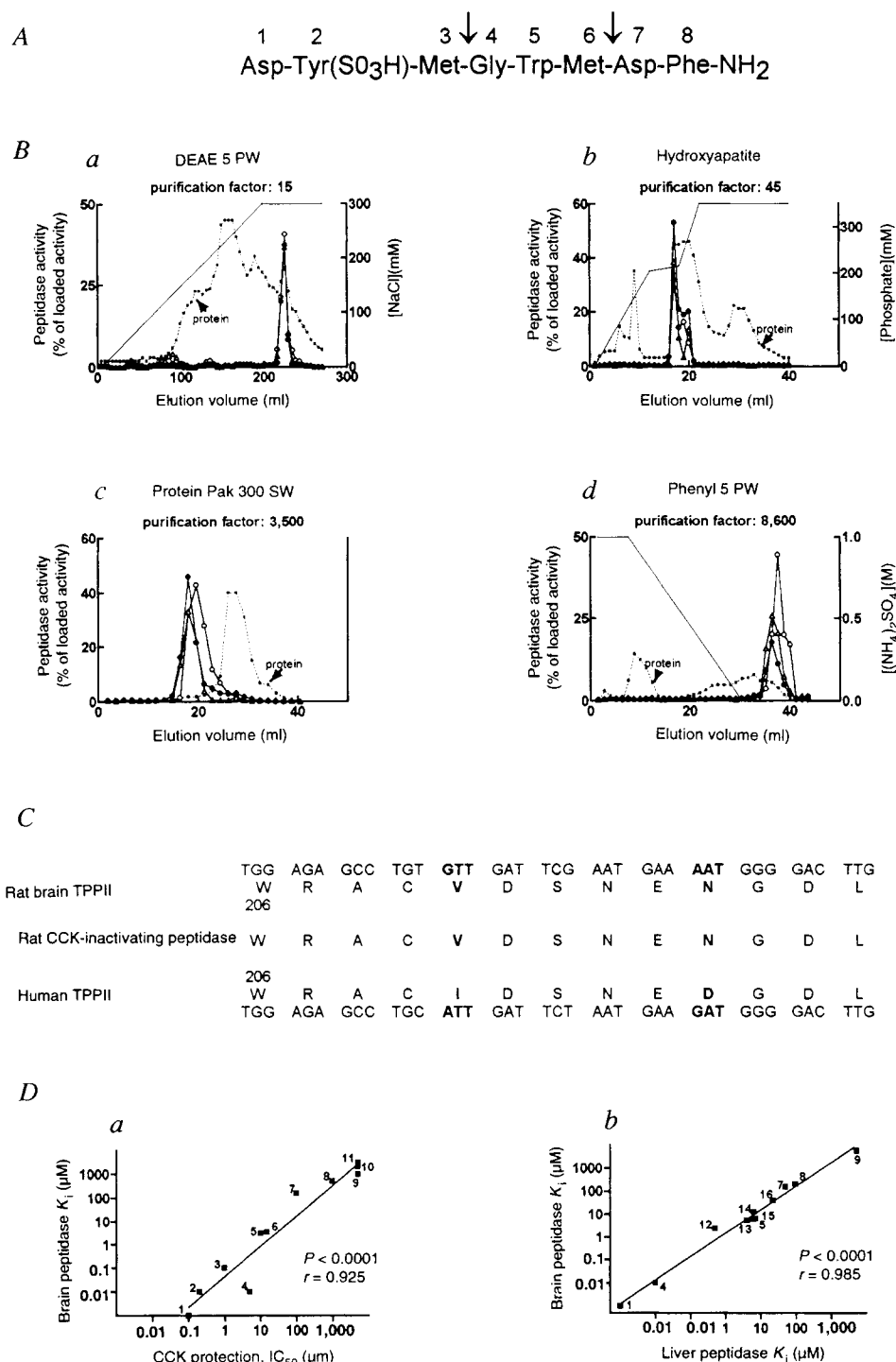
Membrane-associated tripeptidyl peptidase II

Microsequencing of three peptides generated by hydrolysis of the 135K peptide led to amino-acid sequences that were found by computer search in Genbank to display a high degree of homology with sequences 206–218, 857–867 and 1,098–1,112 of human tripeptidyl peptidase II (TPPII) (EC 3.4.14.10)²⁶.

TPPII is a subtilisin-like peptidase that was initially purified from cytosolic extracts of rat liver and human erythrocytes^{27,28} and is found in a variety of tissues; its physiological role was essentially

FIG. 1 Purification and identification of a CCK-inactivating peptidase from cerebral membranes. **A**, Various cleavages (arrows) of endogenous CCK-8 by serine peptidase(s) of brain slices and assay systems used for purification. The responsible peptidase activities were evaluated by measurement of GWM formation from three distinct substrates: GWMDF-NH₂ (CCK-5) for Met6-Asp7 cleavage, DYMGWM for Met3-Gly4 cleavage, and CCK-8 (non-sulphated) for both cleavages together. **B**, Four successive HPLC columns (1–4) were used to purify solubilized peptidase activities forming GWM from CCK-5 (open circles), DYMGWM (open triangles) and CCK-8 (filled circles). The purification factor is the ratio of specific activities of the preparation and the starting material, CCK-8 being the substrate. Similar purification factors were obtained using the other substrates. **C**, Amino-acid sequence of one peptide isolated after hydrolysis of the purified peptidase as compared to that derived from the nucleotide sequences of h-TPPII (ref. 26) and rat TPPII, cloned from a brain cDNA library. **D**, Inhibition constants for the CCK-hydrolysing peptidase activities purified from either brain membranes or liver cytosol, and the peptidase activity responsible for endogenous CCK inactivation in brain slices. K_i values were derived⁴¹ from their IC_{50} values for hydrolysis of 50 μ M AAF-Amc ($K_M = 25 \mu$ M). For IC_{50} values for inhibition of degradation of endogenous CCK immunoreactivity released from depolarized brain slices, see ref. 22. Compounds: 1, Ala-Pro-boroVal pinanediol ester (UCL 1372); 2, Abu-indoline-CO-NH-Bu (butabindide); 3, Abu-Pro-NH-Bu (UCL 1371); 4, Ala-Ala-Phe-CH₂Cl; 5, p-chloromercuribenzene-sulphonate; 6, Ala-ProNH₂; 7, phenylmethane-sulphonyl fluoride; 8, D-Phe-Pro-Arg-CH₂Cl; 9, O-phenanthroline; 10, Tos-Lys-CH₂Cl; 11, Tos-Phe-CH₂Cl; 12, CCK-8; 13, Methoxysuccinyl-Ala-Ala-Pro-boroVal; 14, CCK-8 (non-sulphated); 15, Diprotine A; 16, Ala-Ala-Phe-Amc. **E**, SDS-PAGE of the purified enzyme preparation: silver staining (lane 1) and selective labelling with [³H]DFP (lane 2). **F**, Western blot of CCK-inactivating peptidase from cerebral membranes and liver cytosol. Lanes: 1, crude liver cytosol (20 μ g); 2, purified liver enzyme (30 ng); 3, crude cerebral membranes (20 μ g); 4, purified enzyme from cerebral membranes (40 ng).

METHODS. For enzyme assays during purification, substrates (0.1 mM) were incubated at 37 °C in 50 mM K₂/K phosphate buffer, pH 7.4, containing 0.1 mM bestatin and 0.1 μ M thiorphan; GWM formation was evaluated after 30 min, RIA after derivatization²⁵ (C. Rose *et al.*, submitted). With CCK-5 as substrate, GWM formation was evaluated by HPLC, detection at 280 nm (ref. 24). TPPII activity was evaluated using 50 μ M AAF-Amc²⁸ in buffer containing bestatin, thiorphan and 0.1% Brij 35. For enzyme purification, rat cerebral cortex was homogenized in 10 vol buffer containing 10% glycerol. After centrifugation (500,000g), the washed pellet was the enzyme source, with specific activity 0.1 nmol mg protein⁻¹ min⁻¹, using CCK-8 as substrate. This membrane fraction was solubilized using 1% Brij 35 in the glycerol-containing buffer (60 min at 25 °C) and the resulting solubilise (0.05 nmol mg protein⁻¹ min⁻¹) submitted to HPLC on four successive columns. A DEAE-5PW column (Waters) was eluted at 3 ml min⁻¹ using a linear NaCl gradient (0–300 mM over 1 h) prepared in the glycerol-containing buffer supplemented with 0.025% Brij 35 and 1 mM DTT. A biogel APT column (Biorad) was eluted at 0.5 ml min⁻¹ using a nonlinear K₂/K phosphate gradient (0–350 mM over 45 min). A Protein Pak



300 SW column (Waters) was eluted at 0.4 ml min⁻¹ using the supplemented glycerol-containing buffer. A Phenyl 5PW column (Waters) was eluted using a linear (NH₄)₂SO₄ gradient (1M–0M over 50 min). The same procedure was applied to a supernatant of a rat liver homogenate and led to similar elution profiles and a preparation of 2 μ mol mg protein⁻¹ min⁻¹. Purified brain enzyme was cleaved with CNBr and three microsequenced peptides⁴² displayed 95% identity with sequences of hTPPII (ref. 26). For TPPII cDNA cloning, a Lambda ZAP library (Stratagen) was screened using probes corresponding to nucleotides 38–394 and 3,390–3,770 of hTPPII. A full-length coding sequence corresponding to that of the hTPPII cDNA (89% identity) was obtained from 3 overlapping clones. For [³H]DFP labelling, 1 μ g purified brain enzyme was incubated for 2 h at 37 °C with 10 μ Ci [³H]DFP with or without 0.2 mM Ala-Ala-Phe-CH₂Cl followed by Sephadex G-25 chromatography. For western blots SDS-PAGE⁴³ blots were revealed with the IgG fraction⁴⁴ of polyclonal antibodies raised against the liver enzyme and a secondary ¹²⁵I-labelled anti-rabbit IgG.

TABLE 1 Hydrolysis of natural peptide substrates by TPPII purified from rat brain membranes

Compound	Sequence	Hydrolysis rate		K_M or K_I (μ M)	Specificity constant ($\text{min}^{-1} \text{mg protein}^{-1}$)
		($\mu\text{mol mol}^{-1} \text{ml protein}^{-1}$)	(%)		
CCK-5	GWMDF-NH ₂	16.56	(100)	K_I , 18	0.92
CCK-8 (non-sulphated)	DYMGWMDF-NH ₂	4.00	(24)	K_M , 10	0.40
CCK-8 (sulphated)	DYMGWMDF-NH ₂	2.48	(15)	K_M , 1	2.48
Neurokinin A	HKTDSFVGLM-NH ₂	2.40	(15)	K_I , 40	0.06
Somatostatin (1–14)	SANSPAMAPRERK	2.40	(15)	K_I , 75	0.03
Vasopressin	CYFQNCPRG-NH ₂	2.34	(14)	K_I , >300	<0.01
Somatostatin (15–28)	AGCKNFFWKFTSC	1.12	(6)	K_I , 80	0.01
Dynorphin A	YGGFLRR	0.40	(2)	K_I , 22	0.02
VIP	HSDAVFTDNYTR ...	0.52	(3)	K_I , 88	0.01
CGRP	ACDTATCVTHRL ...	0.55	(3)	K_I , 70	0.01

Hydrolysis rates were evaluated after 3 incubations at 37 °C for 15–120 min (selected to obtain linear rates) of peptides (0.1 mM) in the presence of 50 ng of the purified peptidase preparation in 0.1 ml of 50 mM K₂/K phosphate buffer. The intact peptide was isolated from generated fragments using HPLC on a C₁₈ μ Bondapak column eluted with a CH₃CN gradient (21% to 56% over 30 min) buffered with 10 mM AcONH₄ at pH 4.2 and with absorbance detection at 210 nm. K_I values were evaluated⁴¹ from the IC₅₀ values of compounds against hydrolysis of the artificial substrate AAF-Amc (50 μ M) which displays a K_M of 25 μ M. The following peptides were hydrolysed at rates $\leq 0.2 \mu\text{mol min}^{-1} \text{mg protein}^{-1}$ Boc-CCK-8 (non-sulphated), substance P, rANF (1–28), oxytocin, α -MSH, neurotensin, h-angiotensin-I, (Leu 5 and Met 5)enkephalin, enkephalin octapeptide, dynorphin B, β -endorphin, α -Neo-endorphin, galanin, bradykinin, GRF, GRP, NPY, gastrin (1–16), LHRH. K_M values of CCK-8 were obtained from 30-min incubations, using the GWM RIA, and did not differ significantly from K_I values.

unknown²⁹. The catalytic activity of TPPII, which releases tripeptides from the free N terminus of peptides^{26–29}, is consistent with the two-step cleavage of CCK-8 according to the sequential scheme CCK-8 $\rightarrow$ CCK-5 $\rightarrow$ GWM²². The identification of the CCK-inactivating activity purified from brain membranes as TPPII was confirmed by several observations. First, the purified enzyme had a size on SDS-PAGE (Fig. 1f) consistent with that of TPPII, and it cleaved efficiently Ala-Ala-Phe-amidomethylcoumarin (AAF-Amc) with a similar K_M ²⁸. Second, using the fluorogenic substrate AAF-Amc, peptidase activity was detected in cerebral membranes that was inhibited by a series of compounds with potencies highly correlated with their ability to protect endogenous CCK-8 from degradation²² (Fig. 1d). Third, by applying the HPLC purification process developed for brain membranes, a 135K enzyme that produces GWM from CCK-8 was purified to homogeneity from a soluble extract of liver, a tissue with high TPPII activity²⁷. The enzymes from liver cytosol and cerebral membranes displayed very similar sensitivities to inhibitors, and were indistinguishable from them on western blots performed using antibodies raised against the purified liver enzyme (Fig. 1d, f). Finally, using a probe derived from human TPPII, a cDNA with 89% nucleotide identity was cloned from a rat brain library. It also displays 97% identity with the sequence of

mouse TPPII²⁹. Sequences 680–719, 2,633–2,666 and 3,355–3,400 of the rat brain TPPII encoded the three peptides microsequenced from the GWM-forming enzyme purified from rat brain membranes (Fig. 1c).

The hydropathy profile of brain TPPII does not contain any hydrophobic stretch compatible with a transmembrane domain role, suggesting that it cannot be regarded as a true integral membrane protein. However, alternative splicing, a process that could theoretically generate two isoforms of the enzyme, one cytosolic and the other membrane-associated, has been shown to occur in the enzyme in mouse²⁹ and rat brain (not shown). Several plasma membrane proteins appear to be anchored in the lipid bilayer by a C-terminal glycosyl phosphatidyl inositol (GPI) moiety^{30,31}. A characteristic feature of these proteins is the ability of bacterial phospholipases C to release them in soluble form^{30,32}. When a phospholipase C preparation of *Bacillus cereus* of broad specificity, known to contain a GPI-specific phospholipase C, was applied to purified synaptosomal membranes, a sizeable fraction of TPPII activity (up to 50%; to the same extent as Brij 35) was released in soluble form (Fig. 2). Neither neprilysin nor acetylcholinesterase, which are known to be anchored by transmembrane domains^{33,34}, was released. NaCl (0.5 M) also failed to release TPPII from synaptosomal membranes. Thus it seems that the membrane-bound isoform of TPPII is anchored by a covalent GPI link.

Substrate specificity

Soluble TPPII is regarded as an exopeptidase of broad substrate specificity^{26–28}. Nevertheless, important differences were observed in the rate of hydrolysis of more than 30 neuronal and hormonal peptides by purified TPPII (Table 1). CCK derivatives displayed the highest hydrolysis rate, but cleavage of neurokinin A, somatostatin and vasopressin occurred at only slightly lower rates, and calcitonin gene-related peptide, dynorphin A and vasoactive intestinal peptide occurred at a tenfold lower rate; significant hydrolysis of other neuropeptides could not be detected. CCK-5 displayed the highest hydrolysis rate, which is consistent with data obtained with brain slices^{22–24}, and explains why the rapid formation of GWM from CCK-5 is an efficient inactivation process, as CCK-5 retains some biological activity but GWM does not.

A specificity constant, that is, the ratio of rate over K_M (or K_I), which reflects to some extent the rate of hydrolysis of these neuropeptides at low, presumably physiological concentrations, was calculated (Table 1). Constants obtained for CCK derivatives, particularly CCK-8 in sulphated form, were well above those of other neuropeptides, suggesting that TPPII displays a certain degree of biochemical preference for CCK. However, the

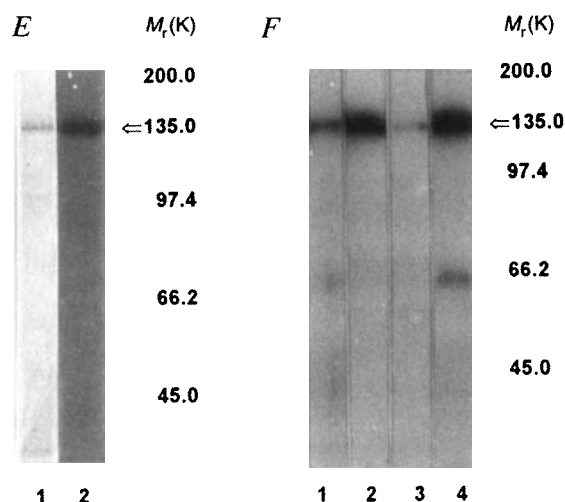
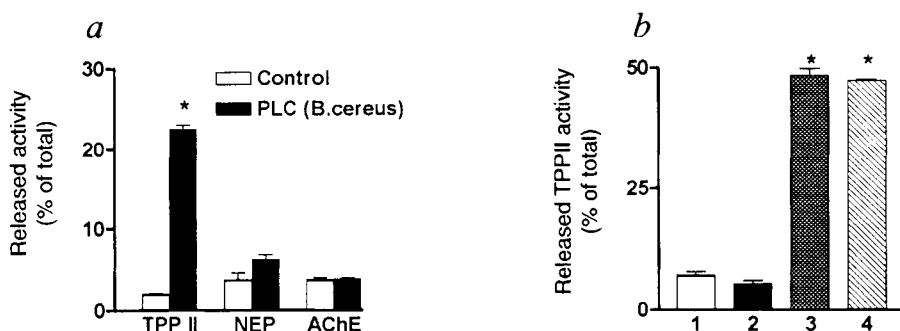


FIG. 2 Phospholipase C-induced release of TPPII activity from synaptosomal membranes of rat brain. *a*, Effects of phospholipase C (PLC) from *B. cereus* (3 U ml^{-1}) on release of TPPII, neprilysin (NEP) and acetylcholinesterase (AChE). * $P < 0.001$ as compared with control. *b*, Effects of 0.5 M NaCl (2), PLC (20 U ml^{-1}) (3), and 1% Brij 35 (4), as compared with controls (1). * $P < 0.001$.

METHODS. Synaptosomal membranes were prepared⁴⁵ corresponding to fractions F and G of the second gradient, and washed twice with phosphate buffer. Incubations were performed at 37°C for 45 min (*a*) or 3 h (*b*) with 0.3 mg membrane protein and PLC from *B. cereus* (type IV, Sigma) in 0.2 ml phosphate buffer. NaCl (0.5 M) or Brij 35 (1% in buffer) were in contact with membranes for 1 h at 37°C or 30 min at 20°C ,



respectively. After centrifugation ($400,000g$), pellets and supernatants were assayed for TPPII (using AAF-Amc; see Fig. 1 legend), neprilysin⁴⁶ and acetylcholinesterase⁴⁷ activities.

hormonal peptide CCK-33 was not significantly inactivated, as shown by RIA²⁰.

Localization

Adequate localization is an important criterion for 'neuropeptidases'³⁵. TPPII catalytic activity, immunoreactivity and gene transcripts were widely distributed (Fig. 3). No precise localizations of TPPII have been reported so far, but our data are consistent with the previous detection of TPPII mRNAs on northern blots of mouse testis, kidney, spleen and liver²⁹.

A variable fraction of TPPII activity was associated with membranes, ranging from 31% in brain to only 4% in liver (Fig. 3a). In agreement, TPPII activity was recently detected in synaptosomal membranes from human brain²¹, whereas in liver it was considered as purely cytosolic²⁷. The dual localization of TPPII is also well illustrated in electron micrographs of pyramidal neurons of rat cerebral cortex showing TPPII immunoreactivity in the cytoplasm, as well as apparently associated with plasma membranes (Fig. 3f). The hydrolysis of CCK-8 by a TPPII-like activity in extensively washed brain slices^{22,24} indicates that the active site of the peptidase is oriented toward the extracellular space; this is consistent with a role in neuropeptide inactivation.

Furthermore, in many areas of the nervous system, the occurrence of TPPII immunoreactivity or mRNAs is consistent with its role in CCK-8 inactivation. This is the case for primary sensory neurons in the nodose ganglion of the vagus nerve, the dorsal root ganglia of the spinal cord, and the trigeminal ganglion. They express CCK_A receptors which activate corresponding C-fibres when stimulated; this is presumably responsible for the satiating and pro-nociceptive effects of CCK^{7,8}. CCK-immunoreactive axons are found in the vicinity of dendrites and/or axons of these sensory neurons, for example in the gastrointestinal tract, or projection areas, such as the nucleus of the solitary tract³⁶, where TPPII mRNAs were also detected. In the ventromedial nucleus of the hypothalamus, the presence of TPPII is consistent with that of CCK, and its potent excitatory actions through CCK_B receptors³⁷. In the cerebral cortex, hippocampal complex and olfactory bulb, the laminar expression of TPPII is also consistent with the presence of CCK neurons and CCK receptors. In testis, the high expression of the peptidase may correspond to that of CCK mRNAs, presumably in sperm cells³⁸. Nevertheless, the localization of CCK immunoreactivity, CCK receptors and the inactivating enzyme do not always strictly overlap, as is often the case with corresponding markers of other neurotransmitters. Such 'mismatches' also suggest that CCK inactivation is not the sole function of TPPII.

Rational design of inhibitors

Serine proteases have similar catalytic centres but different substrate specificities based on the non-covalent interactions of substrates with subsites of the active site surrounding this centre³⁹.

Most inhibitors synthesized so far comprise: (1) a variable peptide sequence ($P_1, P_2, P_3 \dots$) designed to interact non-covalently with the particular subsites (designated S_1, S_2, S_3 according to ref. 40), and which endow them with specificity; and (2) a serine-reactive group (for example, boronate or chloromethylketone), which endows them with potency (affinity), sometimes by covalent interaction with the catalytic centre. In our design of a TPPII inhibitor with therapeutic potential, we deliberately avoided the use of serine-reactive groups, despite the ease it affords for design, to achieve selectivity, easier membrane penetration (for example, higher bio-availability) and lower toxicity. Our strategy has been, as in the case of the rational design of the neprilysin inhibitor thiorphan^{10,11}, to begin by characterizing TPPII subsites using a series of systematically varied di- and tripeptides, some of which are shown in Fig. 4. Although none of the dipeptides stood out enough to give a reasonable lead, tripeptides showed an increase of over two orders of magnitude in potency, for example between Ala-Ala and Ala-Ala-Ala.

Several features of the tripeptide sequence were identified as being responsible for the higher binding: (1) the absolute requirement of a non-substituted ammonium group, as shown by the dramatic loss of potency when alkylated (Fig. 4) or acylated (not shown); and (2) the major contribution of the amide function to affinity provided by P_1 , so that a tripeptide could be simplified into a dipeptide amide, such as Nle-Nle-NH-Me. These results are consistent with TPPII recognizing a primary amino group and two peptide bonds, and cleaving the third peptide bond. In addition, they showed the favourable influence of a Pro residue in P_2 which was retained to enhance *in vivo* metabolic stability and membrane penetration by decreased hydrogen-bonding capacity.

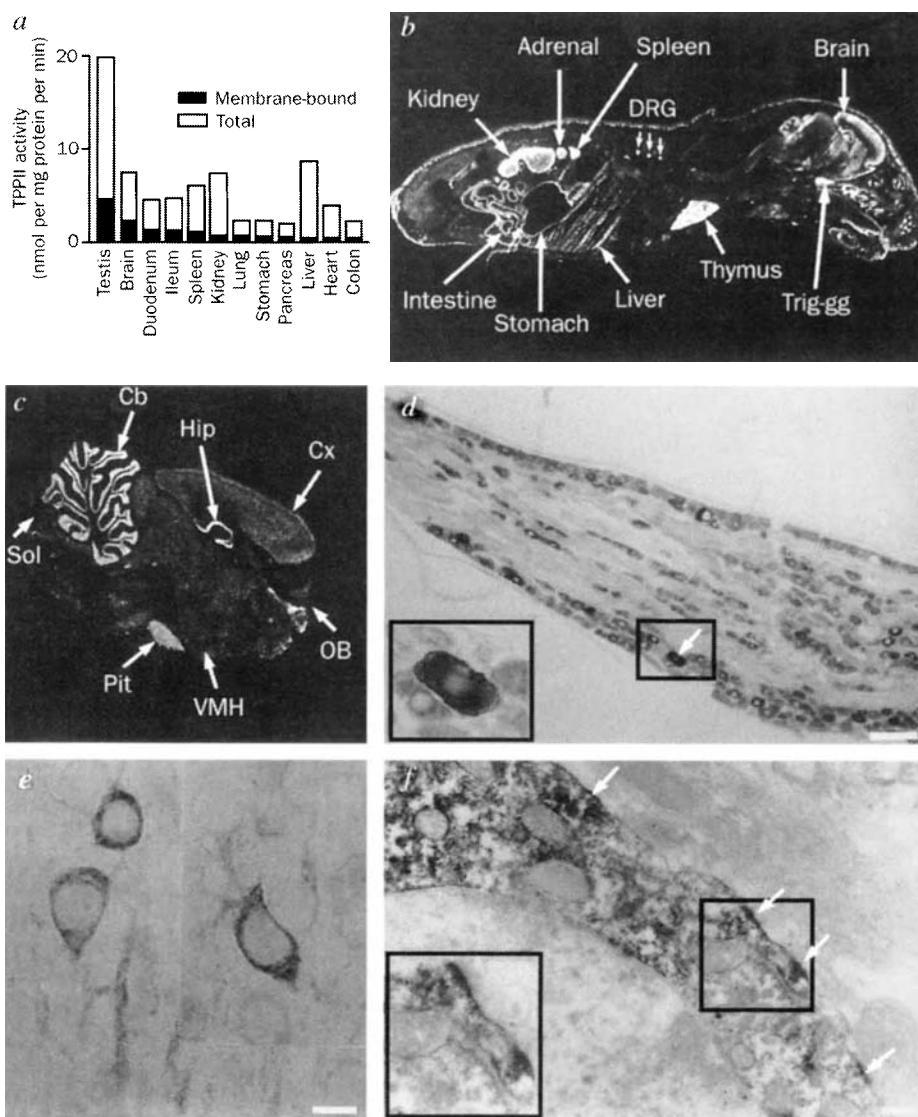
Starting from the sequence P_3 -Pro-NH-R¹, we first optimized P_3 thereby selecting the unnatural amino acid 5- α -aminobutyric acid (Abu). We then optimized R¹ by selecting a butyl chain, because alkyl chains that were either branched or longer than butyl reduced affinity, which suggests that R¹ contributes a hydrophobic binding that is sterically limited but does not act by affecting lipophilicity *per se*. Variants of Pro using other cyclic amino acids were much less active as inhibitors, indicating a high stereochemical dependence and close fit of the optimized compounds to the active site.

In the resulting compound, UCL 1371, which has a K_i of 80 nM, the Pro residue was less bulky than the P_2 residue of the N-terminal tripeptide sequence of CCK-8 (or CCK-5) (Fig. 4c). Hence we tried to locate an accessory aromatic binding region for the inhibitors that would fit the Tyr (or Trp) binding of CCK. Fusion of a benzene ring gave the indoline analogue UCL 1397 (butabindide), which has a K_i of 7 nM; this improved affinity by one order of magnitude and suggests the presence of an aromatic binding pocket in S_2 .

Butabindide was discovered through lead generation followed by progressive optimization of the lead compound, but we now

FIG. 3 Localization of TPPII and gene transcripts in rat tissues and cells. *a*, Distribution of membrane-bound and soluble TPPII activity among rat tissues, which are ranked according to decreasing membrane-bound TPPII activities. *b*, *c*, *In situ* hybridization of TPPII gene transcripts on sagittal sections of newborn-rat whole body and adult brain. Abbreviations: DRG, dorsal root ganglion; trig. gg, trigeminal ganglion; Cb, cerebellum; hip, hippocampus; OB, olfactory bulb; VMH, ventromedial hypothalamic nucleus; Pit, pituitary; Sol, solitary tract nucleus. *d–f*, Immunohistochemical detection of TPPII in neurons of nodose ganglion (*d*), and rat pyramidal cortex layer III at light (*e*) and electron-microscope level (*f*). Arrows indicate the apparent association of immunoreactivity with membranes. Scale bars: 0.1 mm (*d*), 10 μ m (*e*) and 0.4 μ m (*f*).

METHODS. Adult (or 4-day-old in *b*) male Wistar rats were used. For TPPII activity, tissues were dissected and homogenized in 10 vol 50 mM K_2/K buffer at pH 7.4 containing 10% glycerol. TPPII activity was measured in membranes, obtained after 4 centrifugations at 500,000 *g* followed by resuspension in buffer, as well as in the supernatant of the first centrifugation using 0.2 mM AAF-Amc as substrate. Data for membranes were corrected for contamination by the supernatant ($\leq 10\%$), as determined by measurement of lactic dehydrogenase activity. For *in situ* hybridization, a ^{33}P -labelled 276-bp antisense cRNA probe (79–355) was prepared⁴⁸. Sections were exposed to film for 12 days. Controls obtained with a corresponding sense probe showed no significant labelling over background. For immunohistochemistry, rats anaesthetized with Na-pentobarbital were perfused with 4% HCHO; ganglia and brain were dissected out and dipped in the same solution for 2 h. Sections, performed with a cryostat (*d*) or vibratome (*e*, *f*), were incubated in the presence of 0.05% Triton X-100 with the anti-TPPII immunoglobulins (1/5,000) overnight at 4 °C, then with biotinylated goat anti-rabbit immunoglobulins (0.5% Vectastain) for 2 h at room temperature; the reaction was revealed using an Avidin-Biotin Com-



plex (ABC-Elite) and peroxidase visualized using diaminobenzidine. For electron microscopy, dehydrated slides were embedded in Epon resin (TAAB 812) and 0.08- μ m ultrafine sections prepared.

think it could have been designed by removing polar groups (CO_2H , NH , OH , SMe) from the substrate product, DYM, assuming that such groups facilitate the dissociation of hydrolysis products from the enzyme and thereby enhance the catalytic rate. This suggests a general approach to the design of enzyme inhibitors, by the removal of polar groups from the substrate products and increasing affinity through the optimization of hydrophobic binding.

Biological effects of butabindide

We know that butabindide inhibited TPPII competitively (Fig. 5*a*) and selectively because, on a variety of other serine peptidases, it displayed K_i values $>1 \mu M$, in comparison with its low K_i value (7 ± 1 nM) on TPPII. Further, it did not show any significant activity in CCK receptor binding tests ($K_i > 0.1$ mM).

The effect of butabindide on endogenous CCK-8 inactivation in brain was assessed by evaluating the recovery of neuropeptides from depolarized slices of cerebral cortex (Fig. 5*b*). This preparation allows the released neuropeptide to come into contact initially with physiologically relevant peptidases that are presum-

ably located close to the nerve endings from which the peptide originates^{22,23}. The major participation of TPPII in the inactivation process is shown here by the nearly fourfold enhancement of recovery of CCK-8 immunoreactivity provided by butabindide with a potency compatible with selective inhibition of the enzyme. Endogenous GWM immunoreactivity, which had a level in the absence of butabindide that was fourfold higher than that of CCK-8, was progressively suppressed in the presence of butabindide (Fig. 5*b*). These observations strengthen the idea that TPPII is responsible for the major inactivation pathway of endogenous CCK-8 in brain, according to the scheme $CCK-8 \rightarrow CCK-5 \rightarrow GWM$ ^{22,24}.

Biological consequences of TPPII inhibition of butabindide *in vivo* were studied in mice after establishing that it had adequate bio-availability; the 50% inhibitory dose (ID_{50}) for TPPII inhibition in liver and brain was 1.1 ± 0.1 and 6.8 ± 0.5 mg per kg intravenous, respectively.

TPPII could be involved in CCK-8 inactivation in the gastrointestinal tract. Thus the CCK-8-induced delay in stomach emptying was enhanced on pretreatment with butabindide (Fig. 5*c*), an

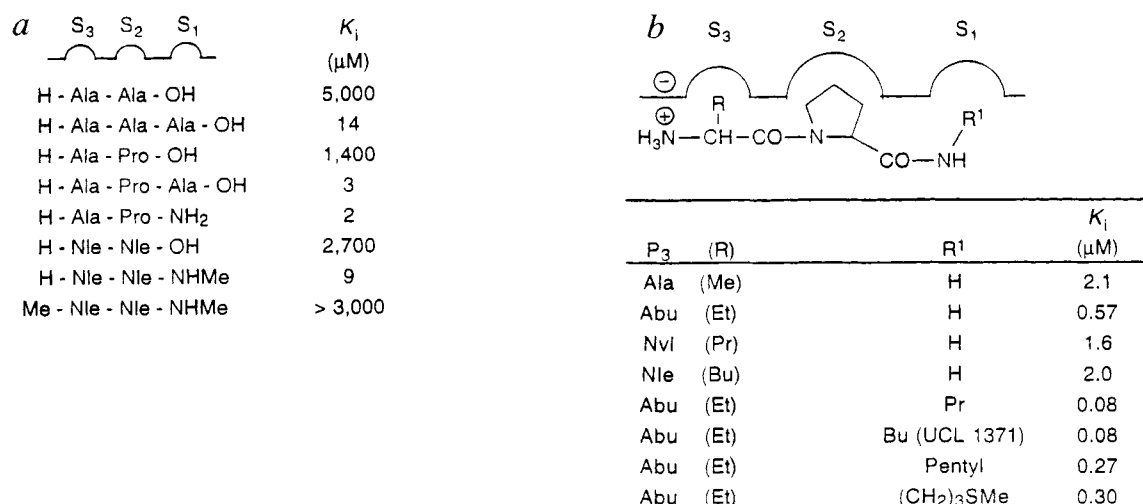
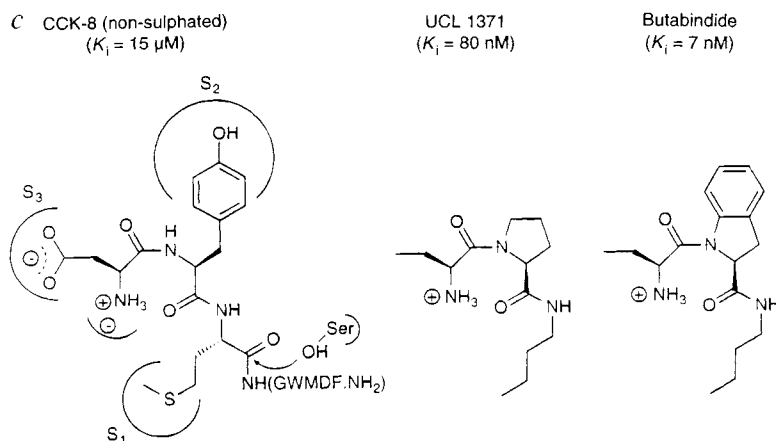


FIG. 4 Stepwise design of a peptoid TPPII inhibitor: from tripeptides to butabindide. K_i values were derived from IC_{50} values of compounds, determined from the inhibition of AAF-Amc hydrolysis of cerebral membranes, and evaluated as described in the Fig. 1 legend. *a*, From tripeptides to dipeptide amides. *b*, Optimization of proline-containing dipeptide amides (P_3 -Pro-NH- R^1). *c*, Optimization of the S_2 -binding residue.



effect blocked by devazepide, a CCK_A receptor antagonist, possibly due to an action on vagal afferent pathways¹. CCK_A receptors, membrane-bound TPPII activity and TPPII mRNAs (Fig. 3) are all detected in rat stomach wall.

Exogenous CCK reduces food intake and elicits other behavioural concomitants of satiation, whereas food intake is increased by systemic administration of CCK_A receptor antagonists⁸. In starved mice, the significant satiating effect of butabindide (Fig. 5*d*), selectivity blocked by devazepide, provides additional support for this idea. Also in rats, butabindide, given just before the lights were turned off, during which time most spontaneous food intake takes place, significantly reduced food intake (not shown). Endogenous CCK-controlling food intake seems to be of neuronal rather than hormonal origin, and acts upon peripheral CCK_A receptors on vagal C-fibres. Several observations are consistent with this view: (1) the satiating effects of butabindide were observed at dosages lower than those required to inhibit TPPII in brain; (2) TPPII immunoreactivity and mRNAs were detected

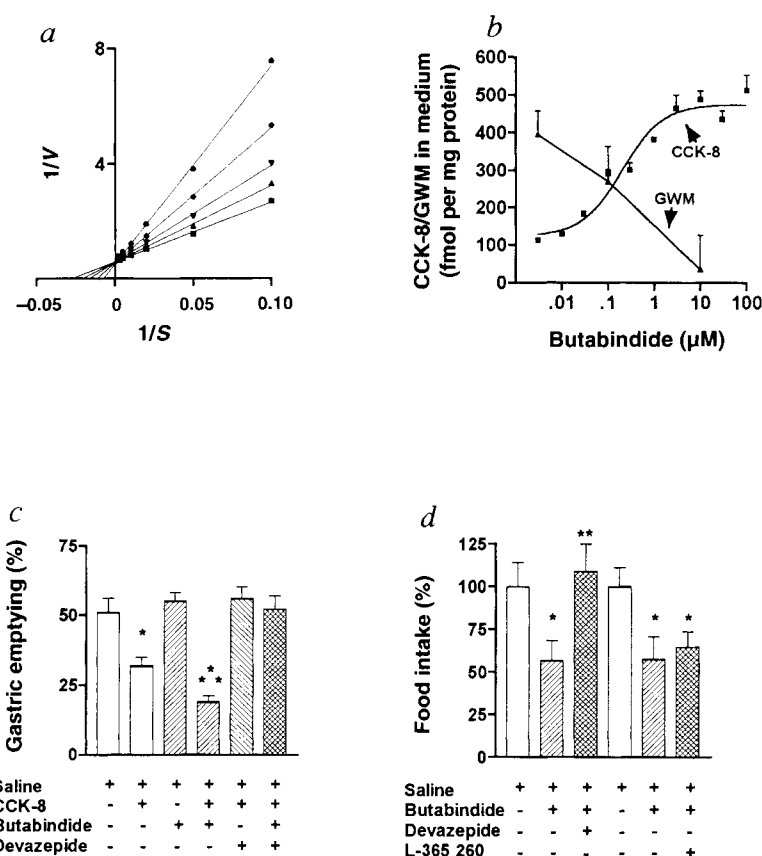
in neurons of the nodose ganglion bearing the CCK_A receptors; and (3) the satiating effect of butabindide is more likely to reflect protection of the neuropeptide CCK-8, an optimal substrate, than of the hormone CCK-33, which is not inactivated by TPPII.

Conclusions

We have established that TPPII fulfills all of the criteria required for it to be considered a 'neuropeptidase'³⁵ responsible for CCK-8 inactivation: (1) in two steps, it rapidly cleaves the neuropeptide into biologically inactive fragments with a reasonably high degree of specificity; (2) it is expressed by CCK-responsive neurons; and (3) its inhibition allows neuronal CCK-8 to escape inactivation and results in CCK-like effects such as satiation in rodents.

Thus TPPII inhibitors should be useful tools to delineate further the functions of CCK neurons, and might become useful drugs for the treatment of disorders such as overeating, problems with gastrointestinal motility, and psychotic syndromes. Nevertheless, additional actions of such inhibitors cannot be excluded,

FIG. 5 Biological effects of butabindide, a TPPII inhibitor. **a**, Competitive inhibition of TPPII activity of cerebral membranes. Membranes of rat cerebral cortex ($0.25 \text{ mg protein ml}^{-1}$) were preincubated for 15 min with or without butabindide (at 0, 3, 6, 12 and 24 nM), and TPPII activity was determined using AAF-Amc ($10\text{--}500 \text{ } \mu\text{M}$) (Fig. 1 legend). **b**, Effects on recovery of CCK-8 and GWM immunoreactivities in the incubation medium of depolarized brain slices. Slices of rat cerebral cortex were depolarized for 5 min by 50 mM K^+ , and peptide levels determined in the supernatant as described²³ except that 0.1 mM bestatin and $0.1 \text{ } \mu\text{M}$ thiorphan were used in studies of GWM (these inhibitors were necessary for protection of the tripeptide but did not modify CCK-8 recovery). Data are differences between levels (fmol per mg tissue protein) in the presence of 50 mM K^+ and 2 mM K^+ . Values without inhibitor: 127 ± 8 for CCK-8, and 395 ± 62 for GWM. Mean $\pm$ s.e.m. of 6 values. **c**, Effects on CCK-induced inhibition of gastric emptying in mice. Mice ($25\text{--}30 \text{ g}$) were food-deprived for 20 h, received orally 0.3 ml solution containing 1 mg phenol red and 0.5% methylcellulose, and were killed 5 min later. Butabindide (10 mg kg^{-1} intravenous, as the oxalate) and devazepide (0.1 mg kg^{-1} intraperitoneal) were administered 30 min before, and CCK-8 (sulphated, $5 \text{ } \mu\text{g kg}^{-1}$ subcutaneous) 5 min before, phenol red. Gastric emptying was calculated as the ratio of phenol red level in small intestine over level in stomach plus small intestine⁴⁹. Mean $\pm$ s.e.m. of 23 values. $*P < 0.01$; $**P < 0.001$ compared with saline and CCK-8, respectively. **d**, Effects on satiety. Mice ($25\text{--}30 \text{ g}$, housed in individual cages, were food-deprived for 20 h (water ad libitum) and as prefeeding received 0.2 ml semi-liquid diet (64% carbohydrate, 10% fat, 25% protein)⁵⁰. After 20 min, a pre-weighted standard food pellet (UAR, France) was presented and weighed after 15, 30 and 60 min. Saline, butabindide (10 mg kg^{-1} , i.v. as the oxalate), devazepide or L365-260, a CCKB receptor antagonist (0.1 mg kg^{-1} , i.p.) were administered 20 min before prefeeding. Results correspond to food intake during



in relationship with the possible participation of the peptidase in the hydrolysis of peptides other than CCK-8. □

Received 13 November 1995; accepted 20 February 1996.

- Dockray, G. J. *Gastrointestinal Endocrinology: Receptors and Post-receptors Mechanisms* (ed. Thompson, J.) 321–332 (Academic, New York, 1990).
- Vanderhaeghen, J. J., Signeau, J. C. & Gepts, W. *Nature* **257**, 604–605 (1975).
- Dockray, G. J. *Nature* **264**, 568–570 (1976).
- Vanderhaeghen, J. J. & Crawley, J. N. *Ann. N.Y. Acad. Sci.* **448**, 1–697 (1984).
- Hughes, J., Dockray, G. J. & Woodruff, G. (eds) *The Neuropeptide Cholecystokinin (CCK)* (Ellis Horwood, Chichester, 1989).
- Hökfelt, T. et al. *Nature* **285**, 476–479 (1980).
- Woodruff, G. & Hughes, J. A. *Rev. Pharmac.* **31**, 469–501 (1991).
- Smith, G. P. & Gibbs, J. *Ann. N.Y. Acad. Sci.* **713**, 236–241 (1994).
- Malfroy, B., Swerts, J. P., Guyon, A., Roques, B. P. & Schwartz, J.-C. *Nature* **276**, 523–526 (1978).
- Llorens-Cortes, C. et al. *Biochem. biophys. Res. Commun.* **96**, 1710–1716 (1980).
- Roques, B. P. et al. *Nature* **288**, 286–288 (1980).
- Patey, G. et al. *Science* **212**, 1153–1155 (1981).
- Baumer, Ph. et al. *Gut* **33**, 753–758 (1992).
- Deschodt-Lanckman, M. et al. *Peptides* **4**, 71–78 (1983).
- Deschodt-Lanckman, M. et al. *Peptides* **5**, 649–651 (1984).
- Matsas, A., Turner, A. J. & Kenny, A. J. *FEBS Lett.* **175**, 124–128 (1984).
- McDermott, J. R. et al. *Neurochem. Int.* **5**, 641–647 (1983).
- Stearns, L. et al. *J. Neurochem.* **45**, 784–790 (1985).
- Durieux, C. et al. *Neuropeptides* **7**, 1–9 (1986).
- Zuzel, K. A., Rose, C. & Schwartz, J.-C. *J. Neurochem.* **15**, 149–158 (1985).
- Wilson, C., Gibson, A. M. & McDermott, J. R. *Neurochem. Res.* **18**, 743–749 (1993).
- Rose, C., Camus, A. & Schwartz, J.-C. *Proc. natn. Acad. Sci. U.S.A.* **85**, 8326–8330 (1988).
- Rose, C., Camus, A. & Schwartz, J.-C. *Neurosci.* **29**, 583–594 (1989).
- Camus, A., Rose, C. & Schwartz, J.-C. *Neurosci.* **29**, 595–602 (1989).
- Llorens-Cortes, C., Gros, C. & Schwartz, J.-C. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6226–6230 (1986).
- Tomkinson, B. & Jonsson, A. K. *Biochemistry* **30**, 168–174 (1991).
- Bälöw, R. M., Ragnarsson, V. & Zetterqvist, Ö. *J. Biol. Chem.* **258**, 11622–11628 (1983).
- Bälöw, R. M. et al. *J. Biol. Chem.* **261**, 2409–2417 (1986).

the 15–60 min period, and are expressed as percent of controls ($140 \pm 20 \text{ mg}$). Mean $\pm$ s.e.m. of 37 values. $*P < 0.05$; $**P < 0.05$, as compared with saline and butabindide, respectively.

- Tomkinson, B. *Biochem. J.* **304**, 517–523 (1994).
- Low, M. G. *FASEB J.* **3**, 1600–1607 (1989).
- Ferguson, M. A. J. A. *Rev. Biochem.* **57**, 285–320 (1988).
- Low, M. G., Stierberg, J., Waneck, G. L., Flavel, R. A. & Kincade, P. W. *J. Immunol. Meth.* **113**, 101–111 (1988).
- Malfroy, B. et al. *Biochem. biophys. Res. Commun.* **114**, 59–66 (1987).
- Massoulié, J., Pezzementi, L., Bon, S., Krejci, F. & Vallette, F. M. *Prog. Neurobiol.* **41**, 31–91 (1993).
- Schwartz, J.-C., Malfroy, B. & De la Baume, S. *Life Sci.* **29**, 1715–1740 (1981).
- Vanderhaeghen, J. J. in *Handbk Chem; Neuroanat.* Vol. 4, Part 1 (eds Bjorklund, A. & Hökfelt, T.) 406–435 (Elsevier, Amsterdam, 1985).
- Hughes, J. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6728–6732 (1990).
- Persson, H. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 6166–6170 (1989).
- Powers, J. C. & Harper, W. *Proteinase Inhibitors* (eds Barrett, A. J. & Salvesen, G.) 55–152 (Elsevier, Amsterdam, 1986).
- Schechter, I. & Berger, A. *Biochem. biophys. Res. Commun.* **27**, 157–162 (1967).
- Cheng, C. & Prusoff, W. H. *Biochem. Pharmac.* **22**, 3099–3108 (1973).
- Tarr, G. E. *Methods of Protein Microcharacterization* (ed. Shively, J. E.) 155–194 (Humana, Clifton, NJ, 1986).
- Laemmli, U. K. *Nature* **227**, 680–685 (1970).
- Danielsen, E. M., Sjöström, H., Noren, O. & Dabelsteen, E. *Biochem. biophys. Acta* **494**, 332–342 (1977).
- Whittaker, V. P. *Ann. N.Y. Acad. Sci.* **137**, 982–998 (1966).
- Gros, B. et al. *J. Pharmac. exp. Ther.* **243**, 666–673 (1987).
- Ellman, G. L., Courtney, K. D., Andres, V. Jr & Featherstone, R. *Biochem. Pharmac.* **7**, 88–95 (1961).
- Sokoloff, P., Gros, B., Martres, M. P., Bouthenet, M. L. & Schwartz, J.-C. *Nature* **347**, 146–151 (1990).
- Lotti, V. J. et al. *J. Pharmac. exp. Ther.* **241**, 103–109 (1987).
- Chariot, J., Nagain, C., Hugonnet, F., Toscan, A. & Rozé, C. *Am. J. Physiol.* **259**, G198–G204 (1990).

ACKNOWLEDGEMENTS. We thank R. L. Henrikson for microsequencing, J. Costentin and A. Blais for pharmacological studies with butabindide, and A. Gaitier for secretarial assistance. The financial support of Upjohn (Kalamazoo) and, at a later date, Bioprojet (Paris), is gratefully acknowledged.

CORRESPONDENCE and requests for materials should be addressed to J.-C.S. The coding sequence corresponding to that of the LTPPII cDNA has been deposited in Genbank, accession no. U50194.

Guide to authors of contributions to *Nature*

Nature is a weekly international journal covering all the sciences. It is intended for an interdisciplinary readership, so all manuscripts should be written clearly and simply. Contributors should pay particular attention to the fact that English is not the first language of many readers. Space is limited and competition severe, so brevity is highly valued.

Nature publishes the following types of article:

Review articles survey recent developments in a topical area of scientific research. Contributors wishing to submit a review should first send a one-page synopsis to the Reviews Coordinator.

- Reviews inform a broad readership about fields where there are rapid, important advances.
- They focus on one topical aspect of a field rather than providing a comprehensive literature survey.
- They can be controversial but in this event should briefly indicate opposing viewpoints. They should not be focused on the author's own work.
- Language should be simple, novel concepts defined and jargon explained.
- Reviews should not be more than 6,000 words and ideally should be shorter. There should be no more than 100 references and ideally half that number. There is no limit to the number of display items or explanatory boxes (used for clarification of technical points or for background material), but reviews do not generally take more than six pages of *Nature*.
- Review articles are often substantially edited by *Nature*'s editors in consultation with the author.

Progress articles review particularly topical and fast-moving fields for a nonspecialist readership. They are similar to Reviews except for the following:

- They do not exceed four pages of *Nature* in length, including display items and references. (One printed page of text is about 1,300 words.)
- They focus on current papers of outstanding interest that are setting new standards in a field.
- Because of their topicality, Progress articles should be written and submitted within a few weeks of *Nature*'s editors expressing an interest in a synopsis.
- Authors may discuss their own work, but should make it clear in the text if they are presenting a personal, rather than a consensus, view.
- Titles are brief (generally a single line) and relatively informal.

Articles are reports of original scientific research.

- They report novel conclusions of broad general scientific interest.
- They represent a substantial advance in scientific understanding of an important problem.
- They should not be longer than 3,000 words, have more than six display items (with captions of fewer than 300 words) or have more than 50 references.
- They have a 'heading' of up to 80 words, often rewritten by *Nature*'s editors, which advertises the content of the paper in general terms. The heading should not contain numbers, abbreviations or measurements unless central to the message of the paper.
- The introduction and summary should be contained in the first two or three paragraphs of the main text.

Letters to *Nature* are short reports focused on a novel, outstanding finding.

- The implications of the main result are of interest to nonspecialists.
- They should not exceed 2 1/2 pages of *Nature*, equivalent to 1,500 words of text with four small display items that have brief captions.
- They have a bold-text first paragraph of not more than 150 words summarizing the rationale for the study, the main result and conclusion. This paragraph is aimed particularly at nonspecialist readers.
- Letters should have fewer than 30 references and no more than four display items. Captions should not exceed 300 words.
- **Articles and Letters** contain a statement at the end of the text: "Correspondence and requests for materials to xxxx". Database accession numbers should be included after this statement.

Preparation of manuscripts

- All manuscripts should be typed, double-spaced, on one side of the paper only.
- Manuscripts should be accompanied by a brief cover letter from the corresponding author, containing a full postal address, telephone and fax number, and e-mail address.
- Five copies of manuscripts and original figures are needed, together with two copies of the covering letter.
- Five copies of relevant related manuscripts in press or submitted for publication elsewhere should be included, clearly marked as such.
- Unless otherwise instructed by the editor handling a manuscript, when sending revised or resubmitted manuscripts, five copies are required, each accompanied by a copy of the authors' response to ref-

erees' and editors' comments on the earlier version.

- Titles should be brief, pertinent and simple, avoiding active verbs, numerical values, abbreviations and punctuation.
- References are numbered sequentially as they appear in the text, tables and figure legends. Only papers that are published or in press should be given numbers: manuscripts submitted or in preparation should be mentioned in the text with a list of authors.
- Reference lists contain only citations to published papers, and do not contain textual material, grant details or acknowledgements.
- Acknowledgements must be brief and appear after the reference list. *Nature* does not publish grant contribution numbers.
- Figures should not be larger than 22 by 28 cm unless unavoidable. They should be marked with the author's name and, where known, the manuscript reference number. One photocopy of the original figures should be provided. All original figures are returned when *Nature* cannot offer to publish a manuscript, but one copy of the manuscript and the photocopy of the figures are kept in confidential files for three months and then destroyed. Detailed instructions for preparation of figures are available from the production department, and digital files of final versions of the figures may be requested.
- Colour figures are welcome but a contribution towards the cost of reproduction is requested. Inability to pay this charge will not prevent *Nature* publishing figures where colour is essential.
- Figures containing protein/nucleotide sequence information should ideally use the three-letter code for amino acids. One column width can accommodate 20 amino acids or 60 base pairs.
- Figure legends should contain fewer than 300 words. They should consist of a brief description of the figure (title, explanation of the parts, explanation of symbols) followed by a telegraphic account of the methods, if appropriate. Multipart figures are discouraged unless the parts are logically connected.
- Tables do not have a 'methods' section. Symbols and abbreviations in the table should be defined immediately below the table, followed by essential descriptive material, all in double-spaced text. Tables should each be presented on a separate sheet of paper.
- As a condition of publication, authors are required to make materials and methods used freely available to academic researchers for their own use. Supporting datasets must be made available at the time of publication either by deposition in the appropriate public database or by distribution on the Internet, together with the relevant accession numbers or site address. In the case of X-ray crystallographic coordinates only, public access may be delayed for up to one year after publication.
- Manuscripts can be submitted to the Editor at *Nature*, Porter's South, Crinan St, London N1 9XW or at *Nature*, 968 National Press Building, Washington, DC 20045-1938. Proofs should be returned by express mail to London.

Other contributions

Nature publishes informal material in each issue, as follows.

- **Correspondence** contributions are short comments on topical issues, anecdotal material or reactions of readers to nontechnical material published in *Nature* (mainly News, Opinion, Book Review or Commentary articles). All are unsolicited; submissions by e-mail to nature@nature.com are preferred (accents on names and places should be clearly indicated). Contributions should be brief (fewer than 500 words).
- **Commentary** articles are informal, editorial comments on topical issues of public interest that have some bearing on scientific research. They are often controversial. Unsolicited manuscripts are considered, but enquiries can be made to the Commentary Editor by e-mail (nature@nature.com) or fax before formal submission as a one-page synopsis or informal letter.
- **News and Views** editorials inform nonspecialist readers about new scientific advances, either as reported in published papers or in the form of meeting reports. Most are commissioned, but proposals can be made to the News and Views Editor in advance of publication of the paper or of the meeting concerned. Authors are not allowed to discuss their own work or work from elsewhere in their laboratory. Detailed guidelines are available on request.
- **Scientific Correspondence** is for informal discussion of topical scientific issues, including material published in *Nature*. Priority is given to contributions of fewer than 500 words. Detailed guidelines are available on request. All contributions are unsolicited.
- **Book Reviews** are all commissioned by the Book Review editor. Unsolicited contributions are not considered.
- *Nature*'s editors will request final copy for Commentary, News and Views, Scientific Correspondence and Book Reviews on diskette or via e-mail.

A more detailed Guide to Authors is available on *Nature*'s www site:
<http://www.nature.com>
<http://www.america.nature.com>

A companion to a quasar at redshift $z = 4.7$

Patrick Petitjean^{*†}, Emmanuel Pécontal[‡],
David Valls-Gabaud[§] & Stéphane Charlot^{*}

^{*} Institut d'Astrophysique de Paris-CNRS, 98 bis, Boulevard Arago,

F-75014 Paris, France

[†] DAEC, URA CNRS 173, Observatoire de Paris-Meudon,

F-92195 Meudon, France

[‡] Centre de Recherche Astronomique de Lyon, CNRS UMR142,

9, Avenue Charles André, F-69561 Saint-Genis-Laval, France

[§] URA CNRS 1280, Observatoire de Strasbourg, 11, rue de l'Université,

F-67000 Strasbourg, France

THERE is a growing consensus that the emergence of quasars at high redshifts is related to the onset of galaxy formation¹, suggesting that the detection of concentrations of gas accompanying such quasars should provide clues about the early history of galaxies. Quasar companions have been recently identified at redshifts up to $z \approx 3$ (refs 2–4). Here we report observations of Lyman- α emission (a tracer of ionized hydrogen) from the companion to a quasar at $z = 4.7$, corresponding to a time when the Universe was less than ten per cent of its present age. We argue that most of the emission arises in a gaseous nebula that has been photoionized by the quasar, but an additional component of continuum light—perhaps quasar light scattered from dust in the companion body, or emission from young stars within the nebula—appears necessary to explain the observations. These observations may be indicative of the first stages in the assembly of galaxy-sized structures.

The bright quasar BR1202 – 0725 was discovered in the automatic plate measuring optical survey for quasars with $z > 4$ (refs 5, 6). The emission redshift is uncertain because Ly α absorption produced by a foreground metal line system is present at a redshift $z = 4.687$ (ref. 7) that is much higher (by $\sim 5,000 \text{ km s}^{-1}$) than the redshift inferred from intermediate-resolution spectroscopy^{7,8} for the C IV emission line of the quasar.

We used the integral field spectrograph TIGER⁹ (Traitement Intégral des Galaxies par l'Etude de leurs Raies) mounted on the 3.60-m Canada–France–Hawaii Telescope during 4–7 February 1995. The instrument samples the field of view through an array of microlenses, each providing a spectrum. Because of windy weather during the observations, only two exposures were useful, corresponding to 2.3 h of integration. The seeing was 0.7 arcsec full-width at half-maximum (FWHM) and the instrumental configuration led to a spatial sampling of 0.61 arcsec in a circular field of view of 10 arcsec diameter centred on the quasar and a spectral sampling of $3.4 \text{ \AA}$. The spectral resolution, measured as the FWHM of an unresolved calibration line, was $10 \text{ \AA}$, and the wavelength range was $5,850\text{--}7,100 \text{ \AA}$, corresponding to a Ly α redshift range of $3.81\text{--}4.84$. The data were reduced using a standard method^{10,11}. The 1σ noise in the spectra obtained through one lens is $3 \times 10^{-19} \text{ erg cm}^{-2} \text{ \AA}^{-1}$ after sky subtraction. We have then reconstructed $60\text{-\AA}$ -wide narrow-band images.

The line of sight towards BR1202 – 0725 contains a known damped Ly α absorption system with an H I column density $N(\text{H I}) \approx 4 \times 10^{20} \text{ cm}^{-2}$ at $z = 4.383$ (refs 7, 12). Such systems are generally thought to arise in proto-galactic disks¹³. They should therefore trace the highest density peaks, and so indicate potential intensive star formation regions. However, we do not detect any emission down to a 3σ limit of $3 \times 10^{-18} \text{ erg cm}^{-2} \text{ \AA}^{-1} \text{ arcsec}^{-2}$ in a region of $41 \times 41 h_{75}^{-1} \text{ kpc}^2$ (where h_{75} is the Hubble constant in units of $75 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and $q_0 = 0.5$) centred on the quasar at $z = 4.383$. The limit reached (which is consistent with previous studies) confirms that damped Ly α systems are not strong Ly α

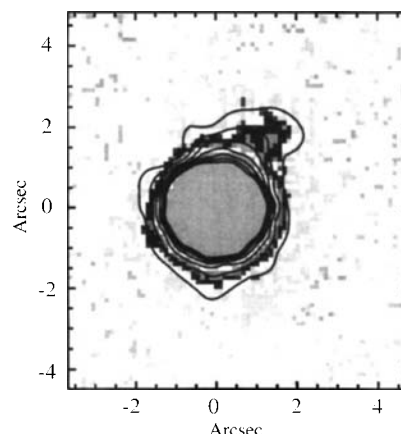


FIG. 1 Isophotes of the emission detected in the wavelength range $6,910\text{--}6,970 \text{ \AA}$ plotted over the r-band image¹⁸. North is up, east is left. A gaussian spatial filter was applied to the merged data cube, leading to a final spatial resolution of 1.2 arcsec (measured as the FWHM of the quasar image in the reconstructed images). The emission detected northwest of the quasar peaks at 2 arcsec from the quasar. Contours of Ly α surface brightness are $0.5, 1.2, 1.9, 2.6, 3.3, 4.0$ and 4.7 times $10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-2}$.

emitters^{14,15} either because of their small but significant dust content¹⁶ or their low duty-cycle of massive star formation¹⁷.

In contrast, the image centred at $6,940 \text{ \AA}$ (see Fig. 1) exhibits a significant (flux $> 3\sigma$) extension to the northwest (region B) with the maximum emission located 2 arcsec away from the quasar (or

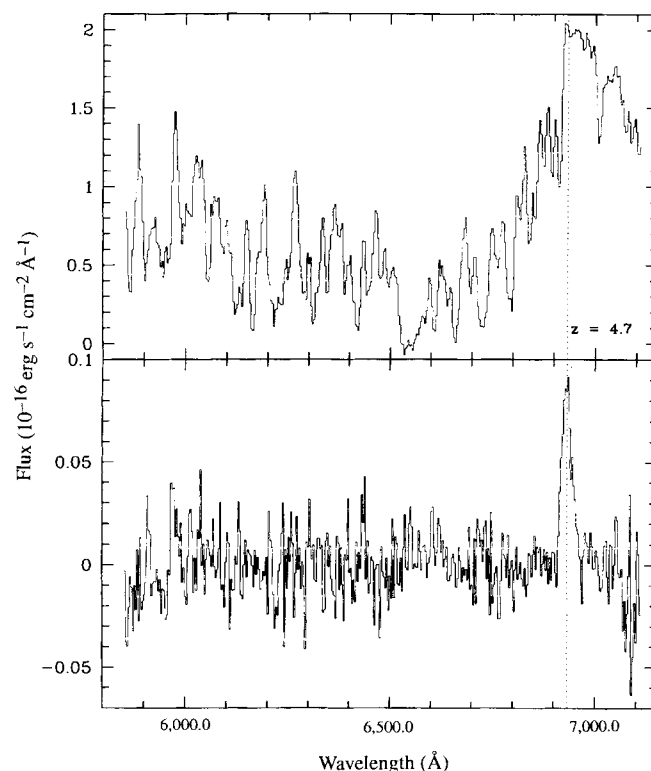


FIG. 2 Spectra of quasar (upper panel) and region B (lower panel). The position of a $z = 4.7$ Ly α line is indicated by a vertical dashed line. The quasar spectrum has been obtained by adding the spectra in all lenses in which its intensity is larger than three times the r.m.s. noise in the reconstructed broad-band image. The quasar spectrum has been subtracted from the spectra from the lenses sampling region B. As the companion is detected only in the Ly α line, the continuum map is an image of the unresolved quasar and thus can be used as the weight map to be applied to remove the quasar contribution everywhere in the field. The spectra from all lenses have then been co-added.

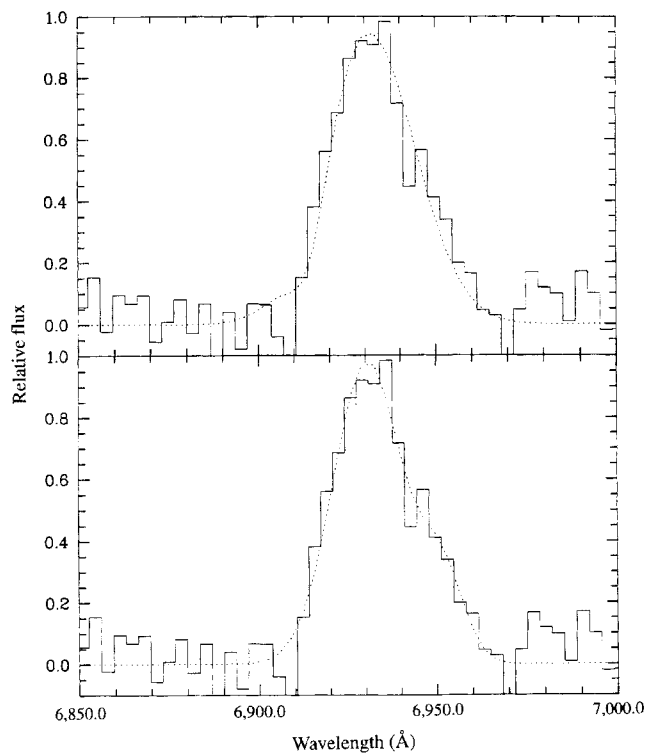


FIG. 3 Model fits of the emission line detected in region B (dashed lines) over the spectrum (solid line). Top panel, model includes a single gaussian profile absorbed in the blue by an intervening absorber. The emission profile peaks at $\lambda = 6,931.5 \text{ \AA}$ and has a width of $1,245 \text{ km s}^{-1}$ (FWHM). The absorption line is centred at $6,914.5 \text{ \AA}$ and has a width of $6 \text{ \AA}$ (FWHM). The reduced χ^2 is 1.04. Bottom panel, model includes two gaussian profiles centred at $6,930.7$ and $6,951.4 \text{ \AA}$ (separation 890 km s^{-1}) and widths (FWHM) of 900 and 410 km s^{-1} , respectively. The reduced χ^2 is 0.96.

$9h_{75}^{-1} \text{ kpc}$) that coincides spatially with an emission excess previously seen in broad-band (V, r, I (ref. 18) and K (ref. 19) bands) and narrow-band filter images²⁰. The contribution of the quasar has been carefully subtracted, and we have co-added the spectra of all lenses in which the extension was detected (Fig. 2). There is a strong line (7σ) centred at $6,932 \text{ \AA}$. The total flux, integrated over the line is $(2.5 \pm 0.6) \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$. Identifying the line with Ly α gives a redshift $z = 4.702$. The emission is spatially resolved and has a maximum northeast/southwest extension of 1.5 arcsec after correction from seeing effects.

As we detect only one line, other potential identifications could be H α at $z = 0.056$, H β at $z = 0.43$ or [O II] $3,727 \text{ \AA}$ at $z = 0.86$. The probability of gravitational amplification is greatest in the latter two cases but population synthesis models indicate that the broad-band observations are inconsistent with these redshifts. In particular, the $r - K$ colour characterizes a steep spectrum that is difficult to reconcile with the lower limit on the B magnitude. Strong absorption systems are present at $z = 1.7544$ and 2.4421 (refs 7, 12) in the quasar spectrum, but these redshifts are ruled out as no emission line is expected at $6,932 \text{ \AA}$.

The detected line coincides in wavelength with the peak of the quasar emission. This is a strong indication that the line is, in fact, Ly α emitted by a quasar companion. Moreover, the blue wing of the line is sharp and coincides almost perfectly with the position of the first Ly α absorption line seen in the quasar spectrum (see Fig. 3). This coincidence can be accounted for if the same absorber extends from the quasar to region B, implying a projected size of at least $9h_{75}^{-1} \text{ kpc}$.

The Ly α flux is 1.7 times larger than the value obtained from previous narrow-band filter imaging²⁰. A thorough check of the photometric quality of our observations using several standard

stars yields an uncertainty of $<20\%$. In addition, the flux determined for the quasar is consistent with published data⁸. The line luminosity is $2.1 \times 10^{43} h_{75}^{-2} \text{ erg s}^{-1}$ and contributes 70% of the r-band emission.

The observed flux is similar to that observed for observed quasar companions at lower redshifts (refs 2, 3, 21) and much larger than the upper limits set by surveys for primeval galaxies in blank fields^{22,23}. This suggests that the presence of the quasar induces a peculiar interaction with the companion, either in the form of enhanced star formation or because the ionizing flux from the quasar contributes to the excitation of the gas. In fact, the ionizing flux of the quasar, estimated from the observed R magnitude and assuming a standard spectral index $\alpha = -1.5$, is

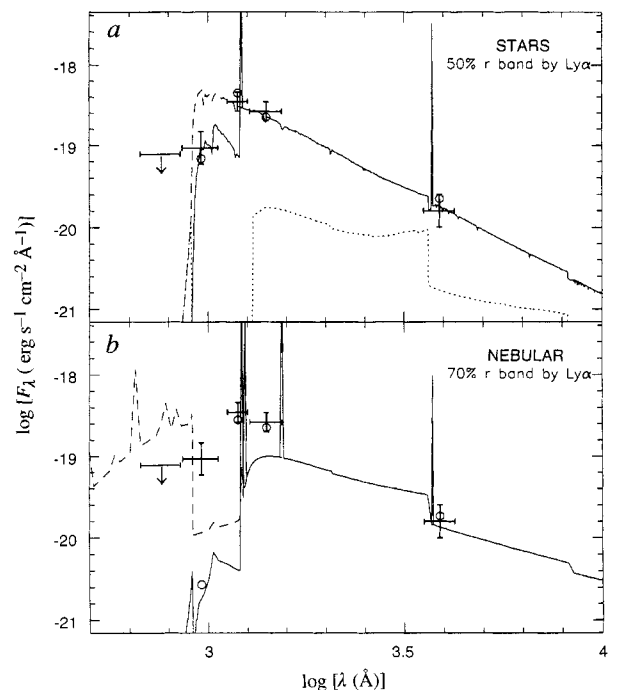


FIG. 4 Model fits to the rest-frame emission properties of region B. Vertical error bars represent uncertainties in the broad-band observations, and horizontal error bars show the filter bandwidths. Open circles represent filter-averaged model fluxes. Also indicated in each panel is the contribution of the Ly α line to the integrated r-band flux. *a*, Spectrum of a $1.1 \times 10^7 \text{ yr}$ -old galaxy with metallicity $0.1Z_{\odot}$, a Salpeter initial mass function with lower and upper cutoffs 0.1 to $100 M_{\odot}$, respectively, and a constant star formation rate of $13 M_{\odot} \text{ yr}^{-1}$, as predicted by recent population synthesis models (G. Bruzual & S.C., manuscript in preparation). Nebular emission was computed²⁷ assuming case B recombination and gas parameters typical of giant H II regions ($n_e = 100 \text{ cm}^{-3}$, $T_e = 10^4 \text{ K}$, $n(\text{He II})/n(\text{H}) = 0.1$, and $n(\text{He III})/n(\text{H}) = 0.01$, where n is the number density, n_e the electron density and T_e blue electron temperature). The dotted line shows the resulting nebular continuum and the solid line shows the sum of the stellar and nebular continua with superimposed Ly α and [O II] emission lines. Blanketing of the radiation to the blue of Ly α , and photoelectric absorption of the (nebular) radiation to the blue of the Lyman limit by foreground H I absorbers at redshifts $z < 4.687$, have been included using approximate analytical formulae²⁸. The dashed line shows the total (stars plus nebular) spectrum in the absence of H I absorption. *b*, Spectrum reflected by a cloud with metallicity $0.1Z_{\odot}$ photoionized by the quasar according to standard photoionization models²⁷. The ionizing spectrum of the quasar has been approximated by a power law $f_{\nu} \propto \nu^{-1.5}$. The electron density, Lyman limit optical depth, and ionization parameter of the model are $n_e = 100 \text{ cm}^{-3}$, $\tau_L = 100$ and $\log U = -1$, respectively, but the quality of the fit depends only weakly on these parameters. To improve the fit, Ly α emission has been reduced arbitrarily by 90%. As in *a*, the dashed and solid lines show the spectrum before and after inclusion of photoelectric absorption and Ly α blanketing by foreground H I clouds.

about $9.3 \times 10^{56} h_{75}^{-2}$ photons s^{-1} . Under case B recombination, the observed Ly α emission of the companion then implies that the fraction of the quasar luminosity intercepted by the gas is $\sim 7 \times 10^{-4}$. If we approximate the cloud as a sphere of diameter $6 h_{75}^{-1}$ kpc, corresponding to a covering factor of 3×10^{-2} , the filling factor of the gas within the cloud is found to be >0.02 . The true value could be even larger if some of the Ly α photons did not escape from the nebula.

The interpretation of the extended emission near the quasar is constrained by the wavelength and equivalent width of the emission line, its contribution to the r-band flux, the upper limit on the B-band emission, and the four VRIK broad-band fluxes. These constraints taken simultaneously rule out the possibility that the emission line is not Ly α at the redshift of the quasar. We consider hereafter two extreme models: a starburst galaxy and a gaseous nebula photoionized by the quasar. Examples of both situations exist at lower redshifts ($z \approx 3$; refs 2–4). Virtually any model galaxy in which the spectrum is dominated by young stars is a reasonable match to the data. Figure 4a shows a typical fit obtained for a model galaxy 1.1×10^7 yr old, with metallicity 0.1 times solar ($0.1Z_{\odot}$) and a constant star formation rate of $13M_{\odot} \text{ yr}^{-1}$. In this case, the onset of star formation would have occurred at $z_i \approx 4.79$ (again for $h_{75} = 1$ and $q_0 = 0.5$). However, in starburst galaxy models the fraction of the r-band flux contributed by Ly α emission is always $<50\%$, that is, much smaller than observed. Therefore such models alone cannot explain the data.

Alternative models of a pure photoionized nebula also present some shortcomings. If the I and K luminosities are assumed to be produced by reprocessing of the quasar's ionizing radiation alone (Fig. 4b shows such a model in which the reflecting cloud has a metallicity of $0.1Z_{\odot}$), the predicted Ly α emission is so strong that acceptable fits of the r-band flux can be obtained only after assuming that a substantial fraction of Ly α photons are removed from the line of sight, because, for example, of orientation effects or absorption by dust¹⁶. Arguing for a small filling factor of the cloud (see above) would not affect this conclusion because this would imply similar changes in both the Ly α and continuum fluxes. The [O II] 3,727 Å flux predicted for the model in Fig. 4b, $5 \times 10^{-18} \text{ erg cm}^{-2} \text{ s}^{-1}$, is consistent with previous limits⁶. The strong predicted N V emission could have been marginally detected (assuming that the cloud has zero metallicity would remove not only N V but also C IV emission, hence worsening the fit of the I-band flux). More important, in addition, is that the model cannot reproduce the observed V magnitude. Hence, although the Ly α emission is most probably a consequence of the excitation by the ionizing flux from the quasar, we conclude that an additional source of continuum is likely to be present.

A possibility is that part of the continuum light from the quasar is scattered in our direction, potentially by dust. In fact, a large amount of dust has been detected in the quasar by continuum emission at millimetre and sub-millimetre wavelengths^{24,25}. New observations suggest that the emission is elongated towards the companion (A. Omont *et al.*, manuscript in preparation), also indicating a large concentration of gas in that direction. We note that, by analogy with radio-galaxies, the ultraviolet continuum light of the companion could result from both scattered light and stellar emission²⁶. The width of the Ly α line ($>1,000 \text{ km s}^{-1}$; Fig. 2) indicates strong disturbances of the kinematics of the gas. Taken together, this suggests that we are witnessing the first stages in the build-up of a galaxy-sized structure. $\square$

10. Ferruit, P. & Pécontal, E. *Astr. Astrophys.* **288**, 65–76 (1994).
11. Bacon, R., Emsellem, E., Monnet, G. & Nieto, J. L. *Astr. Astrophys.* **281**, 691–712 (1994).
12. Lu, L., Sargent, W. L. W., Womble, D. S. & Barlow, T. A. *Astrophys. J. Lett.* (in the press).
13. Wolfe, A. M. *Astrophys. J.* **402**, 411–419 (1993).
14. Pettini, M., Hunstead, R. W., King, D. L. & Smith, L. J. in *QSO Absorption Lines* (ed. Meylan, G.) 55–58 (ESO, Garching, 1995).
15. Lowenthal, J. D. *et al.* *Astrophys. J.* **451**, 484–497 (1995).
16. Charlot, S. & Fall, S. M. *Astrophys. J.* **415**, 580–588 (1993).
17. Valls-Gabaud, D. *Astrophys. J.* **419**, 7–11 (1993).
18. Fontana, A., Cristiani, S., D'Odorico, S., Giallongo, E. & Savaglio, S. *Mon. Not. R. astr. Soc.* (in the press).
19. Djorgovski, S. G. in *Science with VLT* (eds Walsh, J. R. & Danziger, J.) 351–360 (ESO, Garching, 1995).
20. Hu, E. M., McMahon, R. G. & Egami, E. *Astrophys. J. Lett.* (in the press).
21. Hu, E. M., Songaila, A., Cowie, L. L. & Stockton, A. *Astrophys. J.* **368**, 28–39 (1991).
22. Thompson, D., Djorgovski, S. & Trauger, J. *Astr. J.* **110**, 963–981 (1995).
23. Pritchett, C. J. *Publ. astr. Soc. Pacif.* **106**, 1052–1067 (1994).
24. McMahon, R. G., Omont, A., Bergeron, J., Kreysa, E. & Haslam, C. G. T. *Mon. Not. R. astr. Soc.* **267**, L9–L12 (1994).
25. Isaak, K. G., McMahon, R. G., Hills, R. E. & Withington, S. *Mon. Not. R. astr. Soc.* **269**, L28–L32 (1994).
26. McCarthy, P. J. A. *Rev. Astr. Astrophys.* **31**, 639–688 (1993).
27. Ferland, G. J. *Int. Rep.* (Dept of Physics & Astronomy, Univ. Kentucky, Lexington, 1993).
28. Madau, P. *Astrophys. J.* **441**, 18–27 (1995).

ACKNOWLEDGEMENTS. We thank S. D'Odorico, S. Djorgovski and A. Omont for providing us with information before publication. We also thank P. Madau for computing the transmission curve at $z = 4.7$. The observations were made at the Canada-France-Hawaii Telescope, which is operated by CNRS of France, NRC of Canada, and the University of Hawaii.

Orographic control of storm zones on Mars

Jeffery L. Hollingsworth^{*†}, Robert M. Haberle[†],
Jeffrey R. Barnes[‡], Alison F. C. Bridger[§],
James B. Pollack^{||}, Hilda Lee[¶] & James Schaeffer[¶]

^{*} San Jose State University Foundation, San Jose, California 95172, USA

[†] NASA Ames Research Center, MS 245-3, Moffett Field, California 94035, USA

[‡] College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, Oregon 97331, USA

[§] Department of Meteorology, San Jose State University, San Jose, California 95192, USA

[¶] Sterling Software, Inc., Palo Alto, California 94393, USA

Low-pressure cyclones and their accompanying frontal systems in the middle latitudes of the Earth's troposphere develop, travel eastwards and decay preferentially within latitudinally and longitudinally confined geographical regions known as storm zones¹. These zones can be readily identified in the circulation statistics of terrestrial weather systems². Mars, like the Earth, is a rapidly rotating solid planet and has a seasonally varying shallow atmosphere, large-scale orography, and (in a broadly defined context) continental structures³. Although there are also important differences between the two planets, travelling weather systems do exist on Mars³. This raises the question of whether storm zones also occur there, and if so, by what mechanisms. Here we report the results of numerical simulations of global atmospheric circulation patterns on Mars. We find that storm zones can exist during the northern winter, and that continental-scale orography (rather than surface thermal contrasts) is the main factor determining the development of these zones. Storm zones on Mars should play an important role in the martian climate cycle^{4,5}, by influencing the transport of (for example) heat, momentum, water vapour and atmospheric dust^{6–8} towards the poles.

Spacecraft observations leave little doubt that travelling weather systems exist in the martian northern hemisphere. Analyses of surface pressure, temperature and wind data obtained by the Viking Landers have demonstrated clear examples of

Received 16 November 1995; accepted 15 February 1996.

1. Haehnelt, M. & Rees, M. J. *Mon. Not. R. astr. Soc.* **263**, 168–178 (1993).
2. Djorgovski, S., Spinrad, H., McCarthy, P. & Strauss, M. A. *Astrophys. J.* **299**, L1–L5 (1985).
3. Steidel, C. C., Sargent, W. L. W. & Dickinson, M. *Astr. J.* **101**, 1187–1195 (1991).
4. Möller, P. & Warren, S. J. *Astr. Astrophys.* **270**, 43–52 (1993).
5. Irwin, M. J., McMahon, R. G. & Hazard, C. in *The Space Distribution of Quasars* (ed. Crampton, D.) **21**, 117–126 (ASP Conf. Ser., San Francisco, 1991).
6. Pahre, M. A. & Djorgovski, S. G. *Astrophys. J.* **449**, L1–L4 (1995).
7. Wampler, E. J. *et al.* *Astr. Astrophys.* (in the press).
8. Giallongo, E. *et al.* *Astrophys. J.* **425**, L1–L4 (1994).
9. Bacon, R. *et al.* *Astr. Astrophys. Suppl. Ser.* **113**, 347–357 (1995).

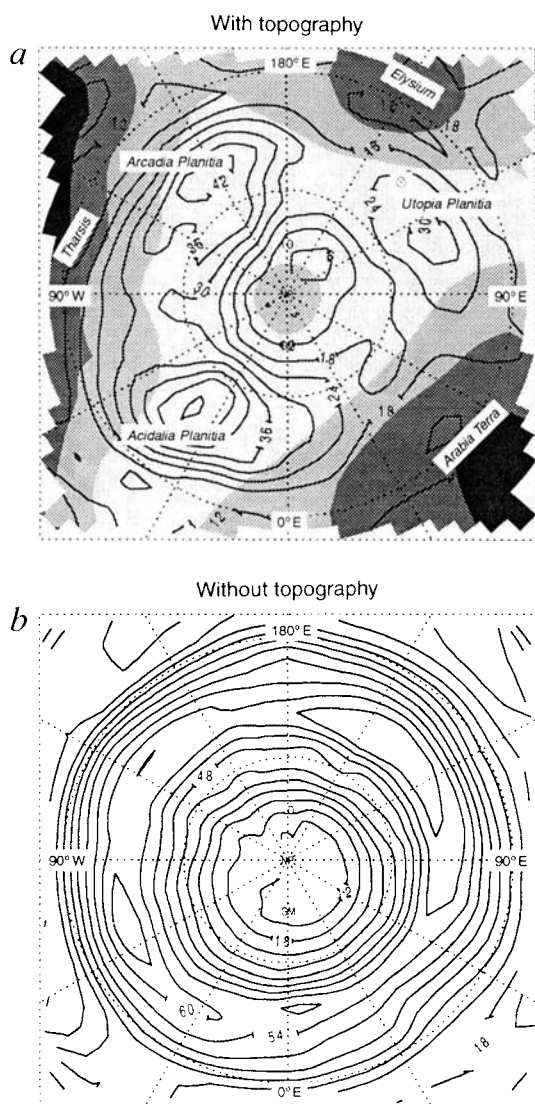


FIG. 1 North polar projections of the transient eddy kinetic energy at the 5-mbar level, $\frac{1}{2}(\overline{u'^2 + v'^2})$ in $\text{m}^2 \text{s}^{-2}$, from a northern winter solstice simulation with the NASA Ames Mars GCM. Here u and v are the eastward (zonal) and northward (meridional) wind components, respectively. The overbar denotes a time average (30 days) and a prime denotes a departure from the average. Growth and decay of these transient disturbances depend on baroclinic and barotropic processes, that is, specific dynamical instabilities within the atmosphere which allow large-scale eddies to amplify and act as the dominant mechanism for the transport of heat, momentum and moisture in winter mid-latitudes^{1,25}. *a*, Simulation with full surface variations in topography, thermal inertia and albedo fields. The topographic height h is depicted by the grey shading; light grey corresponds to $h < 0$ km (that is, the 6.1-hPa mean reference level); medium-light grey, $0 < h < 2$ km; medium grey, $2 < h < 4$ km; and dark grey, $h > 4$ km. Northern protrusions of the high-terrain regions of Tharsis (0.0°N , 110.0°W), Elysium (25.0°N , 150.0°E) and Arabia Terra (20.0°N , 40.0°E) and the low-relief regions (planitia) of Acidalia (45.0°N , 30.0°W), Arcadia (45.0°N , 160°W) and Utopia (48.0°N , 120.0°E) are labelled. *b*, Simulation without topography but with spatially varying thermal inertia and albedo fields. The contour interval in both panels is $6 \text{ m}^2 \text{s}^{-2}$. An integration period of 50 Mars days was made, with time averages and deviations calculated from the last 30 days. Time deviations were low-pass filtered⁹ to remove periods shorter than 1 day. Dotted lines and curves correspond to latitudes and longitudes every 30° and $\otimes$ marks the location of the Viking Lander 2 site (48.0°N , 134.3°E).

transient weather disturbances in the atmosphere of a planet other than Earth. From time spectra of the data, high- and low-pressure systems in middle latitudes with periods of 2–8 days and with east–west scales of $\sim 2,000$ – $4,000$ km have been inferred^{3,9}. But having quite regular periods and near global horizontal scales, martian transient weather systems differ significantly from those on Earth. Furthermore, the present observational data base is too sparse in its spatial and temporal coverage to establish whether storm zones exist on Mars. Thus we examine numerical simulations of the martian atmosphere generated by a three-dimensional global atmospheric circulation model.

The global circulation model used is the NASA Ames Research Center Mars general circulation model (MGCM)¹⁰. The MGCM is a finite-difference model based on the primitive equations of meteorology and includes surface topography using a terrain-following vertical coordinate. Nominally, the model has a resolution of 7.5° latitude $\times$ 9.0° longitude and 13 vertical layers extending from the surface to ~ 47 km (0.067 mbar). The thickness of the layers increases from ~ 500 m at the surface to ~ 5 km at the model top. The topography used is a smoothed version of the

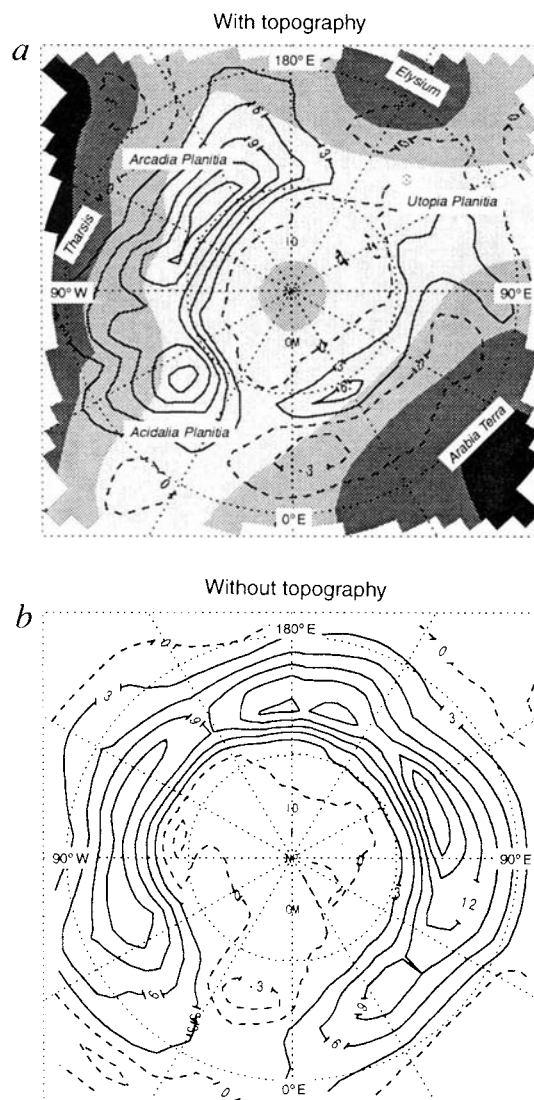


FIG. 2 As in Fig. 1 but for the poleward heat flux $\overline{v'T'}$ (K m s^{-1}). Here, T denotes atmospheric temperature. *a*, Simulation with full surface variations in topography, thermal inertia and albedo fields. *b*, Simulation without topography but with spatially varying thermal inertia and albedo fields. The contour interval in both panels is 3 K m s^{-1} and negative values are shown as dashed lines.

Mars Consortium data set¹¹. Radiative heating rates for the martian tenuous CO₂ atmosphere are calculated allowing for effects of suspended dust on both solar and thermal radiation¹⁰. In the standard configuration, dust is nearly uniformly distributed in space and is 'static' (that is not transported by the circulation). The MGCM has been applied previously to investigate (1) Mars's thermally direct, mean north-south circulation (that is, Hadley cells)¹², (2) its thermally indirect (that is, transient baroclinic eddy or 'weather') components in middle and high latitudes⁸, and (3) its unique seasonal 'condensation flow' associated with the condensation and sublimation of CO₂ within the polar regions¹⁰.

Here we focus on two northern-winter simulations with atmospheric dust loading corresponding to a visible optical depth of 1.0 (ref. 10). One simulation is with full spatial variations in surface topography, albedo and thermal inertia (the 'benchmark' simulation), and the other is with no topography but with spatially varying albedo and thermal inertia. In each simulation, the MGCM is integrated for 50 Mars days from an initial isothermal (170 K) resting state, with the last 30 days used for analysis. Temporal variance and covariance circulation statistics are then constructed as follows. At each MGCM spatial grid point, time series of the prognostic variables are compiled, the time means (30-day averages) are removed, and the temporal deviations are then low-pass filtered⁹ to isolate periods of motions associated with only the transient baroclinic eddies. The time-filtered statistics can then be interpolated onto constant-pressure surfaces and examined for spatially coherent patterns.

Compelling evidence that Mars's mid-latitude weather systems occur prominently in certain regions is indicated in temporal variance and covariance maps from our simulations (Figs 1 and 2). The maps are polar stereographic projections at the 5-mbar level and the top panels correspond to the benchmark case. The transient eddy kinetic energy (that is, total horizontal wind variance) in the benchmark simulation (Fig. 1a) shows that the strongest wind variations occur within the low-relief regions of Acidalia (30.0° W), Utopia (120.0° E) and Arcadia (160.0° W) planitia. This simulated localized variability is associated with developing and decaying weather systems travelling from west to east. The highlands in the eastern hemisphere are clearly more quiescent. Although not shown, similar spatial patterns are revealed in the surface pressure variance. At the MGCM grid point closest to the Viking Lander 2 site, a maximum amplitude of 0.05 mbar² occurs, roughly 50% larger than observations indicate during 'typical' winters⁹.

Without orographic influences, little localization of transient activity is found in the model (Fig. 1b). The kinetic energy becomes highly symmetric in an east-west (that is, zonal) sense. The horizontal regionalization is thus not a result of thermal forcing associated with spatial variations in thermal inertia and albedo fields. Comparing variance magnitudes at this pressure level, the transient eddies are also nearly 25% stronger. But at higher levels, the situation reverses so that the transient eddies in the benchmark case are stronger than those in the 'no-topography' simulation. Orography affects not only the horizontal distribution but also the vertical distribution of transient eddies in a non-uniform way.

An essential aspect of mid-latitude weather disturbances is that they tend to weaken the mean pole-to-equator temperature gradient (that is, baroclinicity) via low-level poleward heat and upper-level momentum fluxes^{13,14}. In the benchmark simulation, the covariance of eddy north-south wind and eddy temperature (Fig. 2a) reveals rather broad maxima on the northern boundaries of high relief and the strongest poleward heat flux occurs in the western hemisphere. In contrast, the highlands of the eastern hemisphere are associated with a flux towards the equator which corresponds to the region of weak kinetic energy (Fig. 1a). In the no-topography simulation (Fig. 2b), a zonally broad band of poleward heat flux can be seen with little localization in the extratropics.

One distinct difference between Earth's storm zones and those

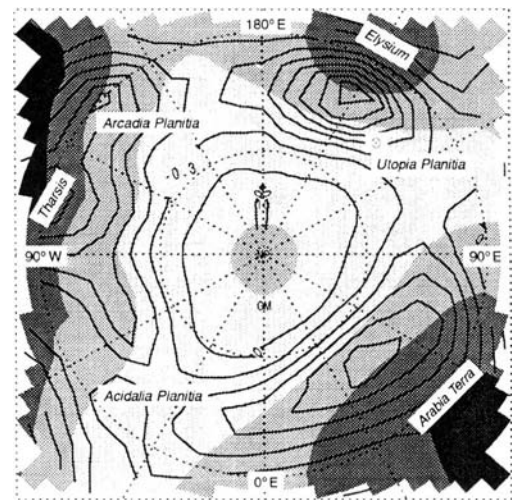


FIG. 3 As Fig. 1 but for the ratio (δ) of the meridional topographic slope to the slope of the time-mean isentropic surfaces in the full MGCM simulation. This ratio is determined using $\delta = (R\bar{T}\theta_\sigma h_\sigma)(afH\bar{\theta}_\sigma)^{-1}$, and its local variations are evaluated on the surface $\sigma = 0.6$ (near the 4.5-mbar level) in the simulation with topography. The contour interval is 0.3. In the above expression, $\theta \equiv T(p_i/p)^{\kappa}$ is the potential temperature, T is temperature, p is pressure, $p_i = 6.1$ mbar which is a reference pressure, h is the surface topography, a is the planetary-mean radius, R is the gas constant and $\kappa \approx 0.26$ for a predominantly CO₂ atmosphere, H is the density scale height, $f = 2\Omega \sin \phi$ is the Coriolis parameter, Ω is the planetary rotation rate, σ is the terrain-following vertical coordinate of the MGCM and ϕ is latitude. An overbar represents a time-averaged quantity (here, as in Fig. 1, this is a 30-day average), and subscripts denote partial derivatives in the latitudinal (ϕ) and vertical (σ) directions. When δ is of the order of 1, topography can significantly diminish growth rates and increase phase speeds of the most unstable baroclinic modes⁷. Values of $|\delta| \approx 1$ are predicted, and suggest regions where instability could be partially suppressed by the strong poleward topographic slopes. The transient eddy kinetic energy maxima (Fig. 1a) appear in the same regions of minima $|\delta|$.

on Mars is the lack of clear separation between the longitudinal scale of the zones and the predominant scales of the transient disturbances. There are two main storm zones in the Northern Hemisphere on Earth, the Atlantic and Pacific zones, just off the east coast of North America and Asia^{1,15}. Earth's northern-winter frontal cyclones have predominant longitudinal scales corresponding to zonal wavenumbers 5–7 (ref. 15). On Mars, three storm zones are apparent within the planitia of the northern hemisphere. The predominant zonal scales of the winter mid-latitude transient disturbances are similar to the physical scales of their terrestrial counterparts, but on the smaller planet Mars these zonal scales are nearly global, ranging between 1 and 4 (refs 8, 9).

Another important difference is that, on Earth, land-ocean thermal contrasts are vital in the maintenance of storm zones^{15,16}. Our simulations, however, indicate that surface thermal effects are weak on Mars and that continental-scale orography is crucial for storm zones on this planet. This control is even apparent in time-mean fields. For example, considerable east-west asymmetry occurs in the time-mean temperature field, with enhanced north-south gradients over the higher terrain continents. The time-mean zonal winds several scale heights above the surface (that is, below 1–2 mbar) shows jet 'surges' in these areas with speed maxima 20–30% higher than the ambient flow.

Large-scale surface variations will generate stationary planetary waves¹⁷ which can have a significant role in a planet's atmospheric circulation. Wave patterns associated with the stationary component will set up mid-latitude east-west variations in the seasonally averaged flow¹⁵ as well as guide the paths of eastward-travelling weather cyclones¹. Furthermore, without some sort of external forcing mechanism and an associated longitudinally varying com-

ponent of the time-mean circulation, transient eddies would not form in geographically fixed regions¹⁸. The stationary circulation can thus directly influence locations of extratropical storm zones. Hollingsworth and Barnes¹⁹ find that the winter extratropics should be characterized by large-amplitude, stationary waves having deep vertical structures. In the presence of these modes, it is quite possible that the transient disturbances which intensify in the vicinity of jet surges are initially 'triggered' from upstream topography (that is, via flow deflection near a topographic barrier)¹⁶.

Sloping north-south topography can have important dynamical consequences on baroclinic growth through slope-induced divergence and through the slope acting on fluid parcel trajectories^{6,20}. Applying an analytical model of baroclinic instability with a uniform horizontal temperature gradient, Blumsack and Gierasch⁶ considered such slope effects for Mars. They found that for negative slopes, the wavelength and growth rates of the fastest-growing modes were reduced and the poleward heat transport by the eddies was suppressed. One useful measure of such effects is the ratio of the meridional topographic slope to the slope of the time-mean isentropic surfaces. In our benchmark simulation (Fig. 3), eddy growth appears partially suppressed within broad regions associated with the northern highlands where topography slopes rapidly downward towards the pole.

Our general circulation model simulations indicate that northern winter storm zones will form on Mars within the mid-latitude planitia. This regionalization is highly controlled by Mars's continental-scale relief and, in contrast to Earth, depends little on the surface thermal contrasts. Observational support of the storm zones may exist in data obtained from the Mariner 9 and Viking spacecraft. A high correlation exists in the northern planitia between high rock abundance and strong near-surface winds, consistent with fine surface material swept free²¹. Aloft, there are indications from Mariner 9 and Viking images of cloud-form patterns during northern winter. 'Wave' and 'streak' clouds frequently occur in middle latitudes near the same longitudes of the storm zones^{22,23}, but the present observational constraints are clearly weak.

Confirmation of the existence of storm zones on Mars awaits systematic spatial and temporal coverage by the orbiting spacecraft planned as part of future Mars²⁴ missions. Full characterization of the storm zones is likely to require the landing of meteorological packages on Mars, in locations that include those regions expected to show extrema in circulation variability. A surface network distributed at several longitudes within the same extratropical latitude would clearly discern structures associated with Mars's travelling weather systems and their seasonal variations; it would also provide *in situ* verification of the storm zones.

Although we have focused here on the northern winter, storm zones during the southern winter will probably differ considerably from those in the north because significant hemispheric asymmetries exist in Mars's orography. We have at present no *in situ* weather observations from the southern hemisphere of Mars. The extent there of transient wave activity, the presence of storm zones, and their modification during highly dusty conditions or during seasonal transitions remains to be investigated. □

Received 5 December 1995; accepted 15 February 1996.

- James, I. N. *Introduction to Circulating Atmospheres* (Cambridge Univ. Press, 1994).
- Wallace, J. M., Lim, G.-H. & Blackmon, M. L. *J. Atmos. Sci.* **45**, 439–462 (1988).
- Zurek, R. W. *et al.* in *Mars* (eds Keiffer, H. H., Jakosky, B. M., Snyder, C. W. & Matthews, M. S.) 835–933 (Univ. Arizona Press, Tucson, 1992).
- Kahn, R., Martin, T. Z., Zurek, R. W. & Lee, S. W. in *Mars* (eds Keiffer, H. H., Jakosky, B. M., Snyder, C. W. & Matthews, M. S.) 1017–1053 (Univ. Arizona Press, Tucson, 1992).
- Jakosky, B. M. & Haberle, R. M. in *Mars* (eds Keiffer, H. H., Jakosky, B. M., Snyder, C. W. & Matthews, M. S.) 969–1016 (Univ. Arizona Press, Tucson, 1992).
- Blumsack, S. L. & Gierasch, P. J. *J. Atmos. Sci.* **29**, 1081–1089 (1972).
- Barnes, J. R. *J. Atmos. Sci.* **41**, 1536–1550 (1984).
- Barnes, J. R. *et al.* *J. geophys. Res.* **98**, 3125–3148 (1993).
- Barnes, J. R. *J. Atmos. Sci.* **37**, 2002–2015 (1980); **38**, 225–234 (1981).
- Pollack, J. B., Haberle, R. M., Schaeffer, J. & Lee, H. J. *geophys. Res.* **95**, 1447–1473 (1990).
- Keiffer, H. H., Davis, P. A. & Soderblom, L. A. in *Proc. Lunar planet. Sci. Conf.* **12**, 1395–1417 (1982).

- Haberle, R. H. *et al.* *J. geophys. Res.* **98**, 3093–3124 (1993).
- Hoskins, B. J. in *Large-Scale Dynamical Processes in the Atmosphere* (eds Hoskins, B. & Pearce, R.) 169–199 (Academic, San Diego, 1983).
- Hoskins, B. J., James, I. N. & White, G. H. *J. Atmos. Sci.* **40**, 1595–1612 (1983).
- Grotjahn, R. *Global Atmospheric Circulations: Observations and Theories* (Oxford Univ. Press, New York, 1993).
- Hoskins, B. J. & Valdes, P. J. *J. Atmos. Sci.* **47**, 1854–1864 (1990).
- Held, I. M. in *Large-Scale Dynamical Processes in the Atmosphere* (eds Hoskins, B. & Pearce, R.) 127–168 (Academic, San Diego, 1983).
- Valdes, P. J. & Hoskins, B. J. *J. Atmos. Sci.* **46**, 2509–2527 (1989).
- Hollingsworth, J. L. & Barnes, J. R. *J. Atmos. Sci.* **53**, 428–448 (1996).
- Mechoso, C. R. *J. Atmos. Sci.* **37**, 1393–1399 (1980).
- Greeley, R., Lancaster, N., Lee, S. & Thomas, P. in *Mars* (eds Keiffer, H. H., Jakosky, B. M., Snyder, C. W. & Matthews, M. S.) 730–766 (Univ. Arizona Press, Tucson, 1992).
- French, R. G., Gierasch, P. J., Popp, B. D. & Yerdon, R. *J. Icarus* **50**, 468–493 (1981).
- Kahn, R. *J. geophys. Res.* **88**, 10189–10209 (1983).
- Space Studies Board *An Integrated Strategy for the Planetary Sciences: 1995–2010* (National Acad. Press, Washington DC, 1994).
- Pierrehumbert, R. T. & Swanson, K. L. A. *Rev. Fluid Mech.* **27**, 419–467 (1995).

ACKNOWLEDGEMENTS. We thank R. E. Young, H. C. Houben and J. R. Murphy for comments, and M. Collins and R. Zurek for reviews of an earlier version of this manuscript. This research was initiated when J.L.H. was a Resident Research Associate of the National Research Council; further support was provided by the NASA Planetary Atmospheres Program.

Dominance of mineral dust in aerosol light-scattering in the North Atlantic trade winds

X. Li, H. Maring, D. Savoie, K. Voss & J. M. Prospero

University of Miami, Rosenstiel School of Marine and Atmospheric Science, 4600 Rickenbacker Causeway, Miami, Florida 33149, USA

ATMOSPHERIC aerosols can affect climate by scattering and absorbing solar radiation^{1–3}. Most recent studies of such effects have focused largely on anthropogenic sulphate aerosols, which are believed to exert a substantial cooling influence². Mineral dust aerosols have been largely ignored, because it was thought that their scattering efficiency and concentrations were too low to have a substantial effect on climate. Here we report measurements of the light-scattering properties of North African dust delivered to Barbados by the North Atlantic trade winds. Although the mass scattering efficiency of the dust is only about a quarter of that of non-seasalt sulphate over the North Atlantic⁵, the annual-mean dust concentration in Barbados trade-wind air is 16 times that of non-seasalt sulphate⁶. The net scattering by mineral dust is therefore about four times that by non-seasalt sulphate aerosols. African mineral dust should therefore be the dominant light-scattering aerosol throughout the tropical and subtropical North Atlantic region. Our observations suggest that mineral dust could be an important climate-forcing agent over this ocean region and in other regions where dust concentrations are high^{7,8}.

Climate models incorporating aerosols suggest that anthropogenic non-seasalt sulphate (n.s.s. SO₄²⁻) aerosols in the Northern Hemisphere scatter as much as 1–2 W m⁻² of incoming solar radiation back to space^{2,4}, yielding a cooling rate comparable to the heating rates attributed to greenhouse gases^{4,9}. As a consequence, climate-related aerosol measurement programs have focused largely on anthropogenic n.s.s. SO₄²⁻.

In many regions, however, mineral dust is a common aerosol component^{7,8}. Large quantities of dust are transported from sources in North Africa across the tropical Atlantic, as documented by measurements made continuously on Barbados (13° 15' N, 59° 30' W) since 1965^{6,10,11}. Satellite imagery shows that dust outbreaks take about a week to reach the Caribbean¹²; the dust clouds typically cover a large area in a latitude band 10–20° wide, reaching to altitudes of 3–6 km (refs 12–14). Dust is the dominant aerosol constituent in this region during the spring, summer and early autumn when dust events are most common^{6,10}.

We measured aerosol physical and chemical properties from 4

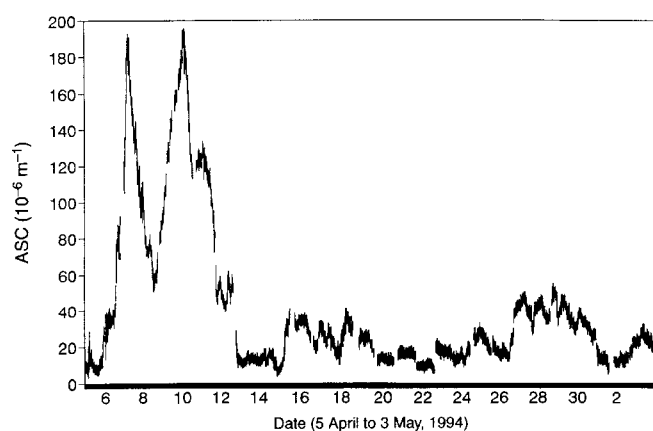


FIG. 1 Time series of ASC_{total} (five-minute averages) measured at Barbados, 5 April to 3 May, 1994 using a Radiance Research Inc. M903 nephelometer at 530 nm.

METHODS. The instrument zero was set daily by flushing with particle-free (filtered) air; it was calibrated by flushing with CO_2 ($ASC = 2.14 \times 10^{-5} m^{-1}$ at $20^\circ C$). Measurements were conducted at the top of a 16-m aluminum tower located on a 30-m-high bluff on the windward coast; the nephelometer sampled directly from the ambient air. To minimize relative humidity (RH) effects on particle growth, the inlet air was heated to ensure that the RH was below 40%; all ASC data in this work apply to these low-RH conditions. A 10- μm -diameter (50% cut-off) impactor was inserted between the heater and the nephelometer to eliminate large particles (principally locally derived sea salt). We periodically alternated nephelometer measurements made with and without the impactor; in most cases the difference in scatter was too small to be detected by the nephelometer, which suggests that the dominant light-scattering aerosols were $<10 \mu m$ diameter. Because light scattering from large particles has a strong forward component and because of the truncated viewing angle of nephelometers, a fraction of the large-particle scatter is not seen by the detector. We estimated the truncation error by Mie scattering calculations using our measured size distributions, and obtained a value of 1.25 which has been incorporated in all ASC values reported here.

April to 3 May 1994. The aerosol scattering coefficient (ASC), an important climate-related property of aerosols², was measured with an integrating nephelometer¹⁵ (results shown in Fig. 1). From 6 to 10 April, dust concentrations (Fig. 2a) were consistently above $65 \mu g m^{-3}$; the value on 9 April, $209 \mu g m^{-3}$, was the highest

ever measured in 30 years at this site¹¹ although daily mean values over $100 \mu g m^{-3}$ are not uncommon. The maximum daily average ASC_{total} ($166 \times 10^{-6} m^{-1}$) coincided with the dust maximum on 9 April (Fig. 2a).

Sun photometer measurements of aerosol optical thickness (AOT), a measure of the attenuation by aerosols of direct solar radiation through the atmosphere, yielded high AOT values through the period, seldom dropping below 0.2 (at 550 nm). The highest AOTs (maximum 0.72) occurred during the large dust event of 8–10 April. The high values of AOT are consistent with past observations^{13,14} that the dust layers are very deep, typically exceeding 3 km.

At times, winds from the central Atlantic brought relatively dust-free air to the site and ASC values dropped sharply (Fig. 2a); nonetheless, substantial quantities of dust were always present. The minimum daily average ASC_{total} ($9.0 \times 10^{-6} m^{-1}$) occurred on 21 April when the dust concentration was $3.3 \mu g m^{-3}$. The concentrations of other aerosol species varied considerably although not nearly as much as that for dust (Fig. 2b). During the very dusty period in early April, the dust concentrations were 30–53 times greater than those of n.s.s. SO_4^{2-} and 9–24 times those of Na^+ ; thus, the ASC_{total} during this period should be dominated by dust scatter.

A scatter plot of daily average ASC_{total} against dust concentration is shown in Fig. 2c. Linear regression yields:

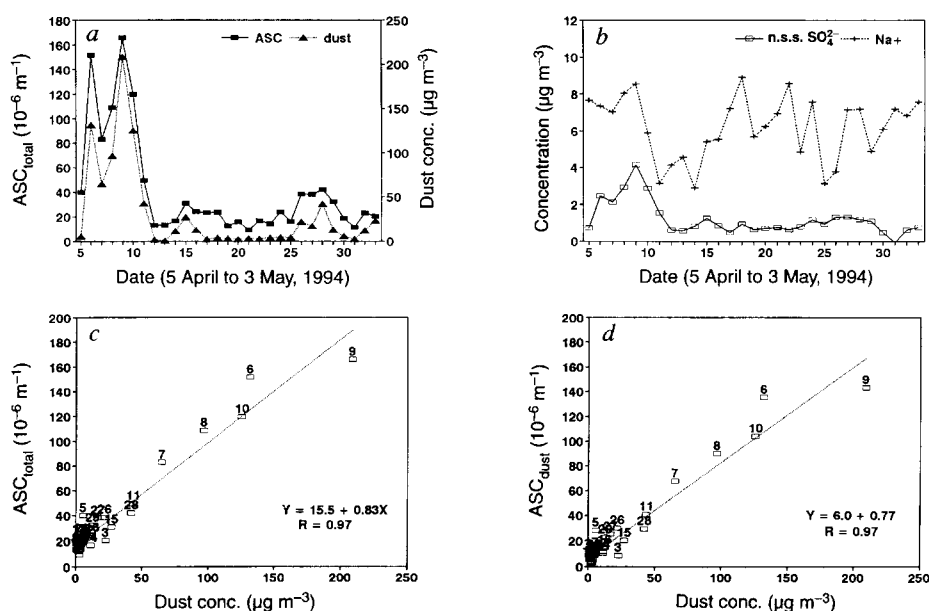
$$ASC_{total} (10^{-6} m^{-1}) = (15.5 \pm 2.3) + (0.83 \pm 0.04) \times [dust] \quad (1)$$

where [dust] is the dust concentration in units of $\mu g m^{-3}$; $r = 0.97$. The slope of the equation provides an estimate of the MSE of dust (MSE_{dust}), $0.83 m^2 g^{-1}$. There is considerable debate about the appropriate definition of MSE, and the procedures that must be used to measure it^{5,16,17}; our approach yields an MSE value that is an ensemble average of the effects of the dominant aerosol component and the minor species that are closely associated with it. The substantial positive intercept in equation (1) suggests that there are other aerosol species that contribute to the value of ASC_{total} independent of dust and associated species. We can estimate the light scatter due to dust alone (ASC_{dust}) by subtracting from ASC_{total} the scattering due to other components:

$$ASC_{dust} = ASC_{total} - \sum (MSE_i \times C_i) \quad (2)$$

where MSE_i and C_i represent the MSE and concentration of aerosol species, i , other than dust. Table 1 shows the mean aerosol

FIG. 2 Relationship between aerosol concentrations and the aerosol scattering coefficients (ASC). a, Time series of daily average of ASC_{total} and mineral dust concentration; b, Time series of daily average n.s.s. SO_4^{2-} and sea salt concentration; c, scatter plot of daily average ASC_{total} against mineral dust concentration; d, scatter plot of daily average ASC_{dust} against mineral dust concentration. In c and d, numbers in the figure refer to the sample date. Daily aerosol samples were collected at a flow rate of $0.8 m^3 min^{-1}$ using Whatman-41 filters which have essentially 100% collection efficiency for marine aerosols²⁸. In Miami, filters are extracted with water; Na^+ is determined by flame atomic absorption and Cl^- , NO_3^- and SO_4^{2-} by suppressed ion chromatography²⁹. Non-seasalt (n.s.s.) SO_4^{2-} is calculated as total SO_4^{2-} minus Na^+ times 0.2517, the SO_4^{2-} : Na^+ mass ratio in sea water. Mineral dust is determined by ashing the extracted filter at $500^\circ C$ for several hours and weighing the residue. To minimize contamination from local sources, pumps were activated only when winds blew directly from the ocean, a condition that was satisfied essentially 100% of the time during the sampling period.



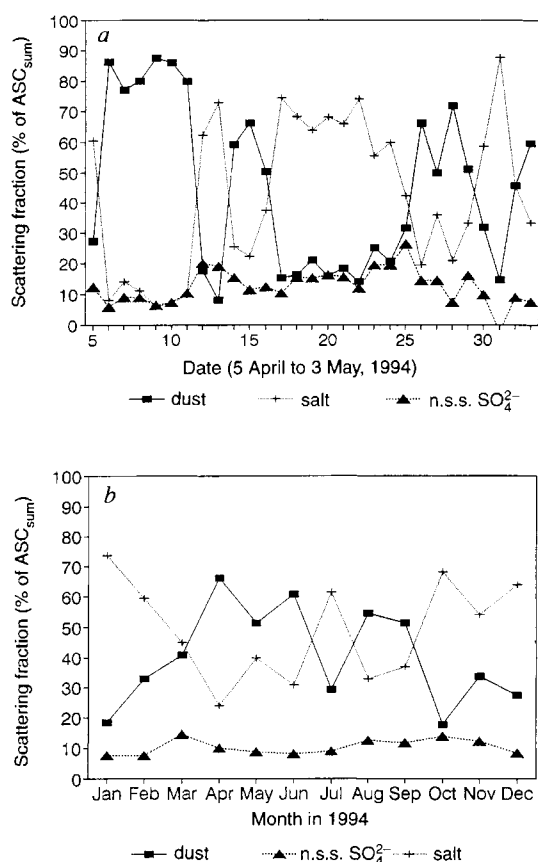


FIG. 3 The relative contribution of mineral dust, sea salt and n.s.s. SO_4^{2-} to the ASC_{sum} at Barbados. *a*, Daily values for the time period of the experiment, 5 April to 3 May, 1994; *b*, monthly mean values calculated for the year 1994.

concentrations of the principal species during the period of the experiment, for all of 1994 and for the period 1984–94. On all timescales, the major aerosol components are mineral dust, n.s.s. SO_4^{2-} and sea salt. Although NO_3^- is a substantial aerosol component, its size distribution follows that of sea salt¹⁸; thus it will not have a significant impact on light scatter relative to sea salt. Ammonium is a minor constituent whose size distribution follows that of n.s.s. SO_4^{2-} ; its scattering effects are apportioned accordingly. Other species are present but they should have a minor effect on light scatter; black carbon concentrations were low, and Saharan dust contains only 3.3% organic matter¹⁹. On this basis in equation (2) we need to account principally for scattering due to n.s.s. SO_4^{2-} and sea salt. We estimated the MSE of sea salt (MSE_{ss}) using data collected when dust concentrations were below $5 \mu\text{g m}^{-3}$; for this sample set the mean seasalt concentration was 7.5 times that of dust. Linear regression analyses of ASC against sea salt yielded estimates of MSE_{ss} that fell in the range 0.3 – $0.5 \text{ m}^2 \text{ g}^{-1}$.

We could not estimate the MSE of n.s.s. SO_4^{2-} ($\text{MSE}_{\text{SO}_4^{2-}}$) at Barbados with the procedure used for dust and sea salt because

sulphate concentrations were always very low compared to dust and sea salt. There are no published estimates of $\text{MSE}_{\text{SO}_4^{2-}}$ for the tropical North Atlantic; as a first approximation we use a value obtained near the Azores⁵, $2.8 \text{ m}^2 \text{ g}^{-1}$, which falls in the low end of the range of previous estimates^{5,16,17} for n.s.s. SO_4^{2-} . A low value for $\text{MSE}_{\text{SO}_4^{2-}}$ is consistent with the observation (based on cascade impactor studies during the experiment) that the size distribution of sulphate is shifted to relatively larger sizes because of the presence of sea salt and mineral dust and the consequent deposition of n.s.s. SO_4^{2-} on the surfaces of these larger particles.

Using equation (2), an intermediate value of $\text{MSE}_{\text{ss}} = 0.4 \text{ m}^2 \text{ g}^{-1}$ and $\text{MSE}_{\text{SO}_4^{2-}} = 2.8 \text{ m}^2 \text{ g}^{-1}$, we calculated ASC_{dust} for the daily samples; a scatter plot of ASC_{dust} against dust concentration is shown in Fig. 2*d*. The regression equation of the line is:

$$\text{ASC}_{\text{dust}} (10^{-6} \text{ m}^{-1}) = (6.0 \pm 2.2) + (0.77 \pm 0.04) \times [\text{dust}] \quad (3)$$

with [dust] in units $\mu\text{g m}^{-3}$. The slope of the equation yields the MSE_{dust} , $0.77 \text{ m}^2 \text{ g}^{-1}$. Because of the extremely high dust concentrations, the slope is relatively insensitive to the values of $\text{MSE}_{\text{SO}_4^{2-}}$ and MSE_{ss} ; a matrix of test values with MSE_{ss} ranging from 0.3 to $0.5 \text{ m}^2 \text{ g}^{-1}$ and $\text{MSE}_{\text{SO}_4^{2-}}$, 2– $5 \text{ m}^2 \text{ g}^{-1}$, yielded a maximum range of 0.73 – $0.79 \text{ m}^2 \text{ g}^{-1}$ for MSE_{dust} . The highest test value for $\text{MSE}_{\text{SO}_4^{2-}}$ ($5 \text{ m}^2 \text{ g}^{-1}$) is the same as that used for the global mean by Charlson *et al.*².

There may be other aerosol scatterers present during dusty periods but their effects, if any, are masked by the extremely high concentrations of dust and, thus, they will have little impact on the derived MSE for dust. Our results are consistent with those obtained in an extensive study of dust optical properties in the western United States by White *et al.*²⁰, to our knowledge the only other study that directly measures dust scatter using similar techniques.

We can estimate the relative contributions of dust, n.s.s. SO_4^{2-} and sea salt and associated aerosol components to aerosol light scatter at Barbados in the boundary layer (reduced to relative humidity $\leq 40\%$) with the equation:

$$\text{ASC}_{\text{sum}} (10^{-6} \text{ m}^{-1}) = 0.4[\text{sea salt}] + 2.8[\text{n.s.s. SO}_4^{2-}] + 0.77[\text{dust}] \quad (4)$$

where ASC_{sum} represents the sum effect of the scattering of the principal species in square brackets (units, $\mu\text{g m}^{-3}$). On a daily basis for April 1994 (Fig. 3*a*), dust was the dominant scatterer during 15 of the 29 days of the experiment. During days when dust was about $10 \mu\text{g m}^{-3}$ or less, sea salt was the dominant scatterer, with a comparable contribution from sulfate and mineral dust.

Studies carried out by us over the past 30 years at Barbados show that the chemical and physical properties (including size distribution and mineralogy) of dust are relatively constant during the year except in the winter when dust concentrations are generally low. On this assumption, we estimate the relative scatter for dust, n.s.s. SO_4^{2-} and sea salt for the year 1994 using monthly mean concentrations for these species (Fig. 3*b*). The contribution from n.s.s. SO_4^{2-} was relatively constant throughout 1994 ($\sim 10\%$ of ASC_{sum}) while that from dust was highly variable, from about 60% during the spring and summer to 20% during the winter. On the basis of the ten-year mean values (Table 1) of aerosol concentrations, the mean scattering contributions to ASC_{sum} from dust, sea salt and n.s.s. SO_4^{2-} are 56%, 33% and 11%, respectively.

Although sea salt is an important light-scattering aerosol, its concentration drops sharply above the marine boundary layer which is typically at an altitude of 1–1.5 km in the western tropical Atlantic; in contrast, continentally derived n.s.s. SO_4^{2-} and dust aerosols in the trade winds are carried in a layer that typically extends to 3–6 km altitude^{13,14,21}. Thus, averaged over the atmospheric column, the net scattering effect of sea salt will be substantially reduced relative to dust and n.s.s. SO_4^{2-} .

TABLE 1 Mean aerosol concentrations at Barbados

Averaging period	dust ($\mu\text{g m}^{-3}$)	n.s.s. SO_4^{2-} ($\mu\text{g m}^{-3}$)	Sea salt ($\mu\text{g m}^{-3}$)	NO_3^- ($\mu\text{g m}^{-3}$)	NH_4^+ ($\mu\text{g m}^{-3}$)
April 1994	30.90	1.19	20.18	0.79	0.13
1994 total	11.16	0.69	21.11	0.62	0.12
1984–94	12.89	0.78	16.50	0.53	0.11

On a global scale, mineral dust is one of the major aerosol components in the atmosphere with emission rates estimated^{7,22,23} to be in the range of about 1,000–3,000 Tgyr⁻¹ while that for n.s.s. SO₄⁻ aerosol is ~250 Tgyr⁻¹ of which 140 Tgyr⁻¹ is from anthropogenic sources²³. Although it is not presently possible to determine the role of dust in light scatter on a global scale on the basis of dust mass emission estimates because of the large uncertainties regarding the variability in the physical properties of the dust, it is clear that dust has the potential to play an important role in radiative processes and climate²⁴.

Although mineral dust is clearly a natural product, the rates of emission of this dust seem to have greatly increased²⁵ owing to drought compounded by land-use practices²⁵. Measurements of mineral dust made on Barbados since 1965^{10,11} show that dust concentrations in the trade winds are correlated with rainfall deficits in the sub-Saharan region. Indeed, the large dust episode in early April 1994 is clearly traceable to the coast of sub-Saharan Africa in satellite (Meteosat and Goes) imagery. Thus, a substantial fraction of the mobilized dust in the atmosphere might properly be considered to be an anthropogenic pollutant. These various factors suggest that the radiative effects of mineral dust could be important in climate processes, and that the effects of dust could complicate the attempts to assess human effects on climate^{24,27}. □

Received 22 December 1995; accepted 11 March 1996.

1. Shaw, G. E. *Atmos. Environ.* **21**, 985–986 (1987).
2. Charlson, R. J., Langner, J., Rodhe, H., Leovy, C. B. & Warren, S. G. *Tellus* **43A**, 152–163 (1991).
3. Hobbs, P. V. in *Aerosol-Cloud-Climate Interactions* (ed. Hobbs, P. V.) 33–73 (Academic, San Diego, 1993).

4. Charlson, R. et al. *Science* **255**, 423–430 (1992).
5. Hegg, D. A., Ferek, R. J. & Hobbs, P. V. *J. geophys. Res.* **99**, 25951–25954 (1994).
6. Savoie, D. L. et al. *J. geophys. Res.* **97**, 11575–11589 (1992).
7. Duce, R. A. et al. *Glob. biogeochem. Cycles* **5**, 193–259 (1991).
8. Prospero, J. M. in *The Oceanic Lithosphere* (ed. Emiliani, C.) 801–974 (Wiley Interscience, New York, 1981).
9. Charlson, R. J. & Wigley, T. M. L. *Scient. Am.* **270**, 48–57 (1994).
10. Prospero, J. M. & Nees, R. T. *Nature* **320**, 735–738 (1986).
11. Prospero, J. M., Savoie, D. L., Arimoto, R. & Huang, F. (abstr.) *Eos* **74**(43), 146 (1993).
12. Ott, S.-T., Ott, A., Martin, D. W. & Young, J. A. *Mon. Weath. Rev.* **119**, 1832–1850 (1991).
13. Prospero, J. M. & Carlson, T. N. *J. geophys. Res.* **77**, 5255–5265 (1972).
14. Carlson, T. N. & Prospero, J. M. *J. appl. Met.* **11**, 283–297 (1972).
15. Ruby, M. G. & Waggoner, A. P. *Environ. Sci. Technol.* **15**, 109–113 (1981).
16. Anderson, T. L., Charlson, R. J., White, W. H. & McMurry, P. H. *J. geophys. Res.* **99**, 25947–25949 (1994).
17. Quinn, P. K., Marshall, S. F., Bates, T. S., Covert, D. S. & Kapustin, V. N. *J. geophys. Res.* **100**, 8977–8991 (1995).
18. Savoie, D. L. & Prospero, J. M. *Geophys. Res. Lett.* **9**, 1207–1210 (1982).
19. Lepple, F. K. & Brine, C. J. *J. geophys. Res.* **81**, 1141–1147 (1976).
20. White, W. H., Macias, E. S., Ninenger, R. C. & Schorran, D. *Atmos. Environ.* **28**, 909–921 (1994).
21. Talbot, R. W. et al. *J. geophys. Res.* **91**, 5173–5182 (1986).
22. Duce, R. A. in *Dahlem Workshop on Aerosol Forcing of Climate* (eds Charlson, R. J. & Heintzenberg, J.) 43–72 (Wiley, Chichester, 1995).
23. Andreae, M. O. in *Future Climates of the World* (ed. Henderson-Sellers, A.) 341–392 (Elsevier, Amsterdam, 1995).
24. Tegen, I., Lacis, A. A. & Fung, I. *Nature* **380**, 419–422 (1996).
25. Glantz, M. H. in *Drought Follows the Plow* (ed. Glantz, M. H.) 33–43 (Cambridge Univ. Press, 1994).
26. Tegen, I. & Fung, I. *J. geophys. Res.* **100**, 18707–18726 (1995).
27. Lacis, A. A. & Mishchenko, M. I. in *Dahlem Workshop on Aerosol Forcing of Climate* (eds Charlson, R. J. & Heintzenberg, J.) 11–42 (Wiley, Chichester, 1995).
28. Pszenny, A., Fischer, C., Mendez, A. & Zetwo, M. *Atmos. Environ.* **27A**, 282–284 (1993).
29. Savoie, D. L., Prospero, J. M. & Saltzman, E. S. *J. geophys. Res.* **94**, 5069–5080 (1989).

ACKNOWLEDGEMENTS. We thank L. Custals and F. Huang for performing chemical analyses; C. Pillaris for assistance in nephelometer truncation error calculations; M. Izaguirre, S. Machani, C. McCormick, C. Shea and T. Snowdon for technical and operational support; and Mrs. E. Manning, Barbados, for use of the site at Ragged Point. This work was carried out as a part of the Atmosphere/Ocean Chemistry Experiment, supported by the National Science Foundation, with additional support from the NOAA Climate and Global Change Program, and the Environmental Protection Agency.

The influence on climate forcing of mineral aerosols from disturbed soils

Ina Tegen^{*†‡}, Andrew A. Lacis[†] & Inez Fung^{†‡}

^{*} Department of Applied Physics, Columbia University, 2880 Broadway, New York, New York 10025, USA

[†] NASA Goddard Institute for Space Studies, 2880 Broadway, New York, New York 10025, USA

[‡] School of Earth and Ocean Sciences, University of Victoria, PO Box 1700, Victoria, British Columbia V8W 2Y2, Canada

AEROSOLS influence the global radiation budget¹, and so changes in the atmospheric aerosol load due to either natural causes or human activity will contribute to climate change². A large fraction of the mass of tropospheric aerosol is wind-blown mineral dust, and its contribution to radiative forcing can be locally significant^{3,22}. Model calculations indicate that 50 ± 20% of the total atmospheric dust mass originates from disturbed soils⁴ (those affected by cultivation, deforestation, erosion, and frequent shifts in vegetation due to droughts and rains). Here, using a radiative transfer model embedded in a general circulation model, we find that dust from disturbed soils causes a decrease of the net surface radiation forcing of about 1 W m⁻², accompanied by increased atmospheric heating that may be a significant forcing of atmospheric dynamics. These findings suggest that mineral dust from disturbed soils needs to be included among the climate forcing factors that are influenced by human activities.

Global dust distributions were calculated with a global model of the dust cycle^{4,5} where dust emissions for eight particle size classes are parametrized in terms of soil moisture, surface wind speed, soil texture, vegetation and soil surface conditions. Each size class is transported in the Goddard Institute for Space Studies (GISS)

three-dimensional atmospheric tracer transport model^{6,7}, and removed from the atmosphere at different rates by gravitational settling, turbulent mixing and rain-out. The model-derived global dust source strength is 1.500 ± 700 Mt yr⁻¹ for particles <10 μm radius. Dust optical thicknesses are calculated from the modelled spatial and temporal distributions of dust mass loading and size distributions⁸.

Retrievals of total atmospheric optical depths over the oceans from the Advanced Very High Resolution Radiometer (AVHRR) on NOAA orbiting satellites are used to constrain the modelled source hypotheses⁴. AVHRR features such as the seasonal shift of the Saharan/Sahelian dust plume and the relatively small contribution of dust from Australia cannot be reproduced if the dust aerosol is assumed to originate only from natural sources. They can be best explained with a model where 50 ± 20% of the total atmospheric dust load is from disturbed soils that are affected by human activities or shifts in climate conditions. The resulting annual global-mean optical thickness of the total dust load is 0.034, of which disturbed soils contribute 0.017.

Radiative properties (extinction efficiency, single scattering albedo, and asymmetry parameter) of aerosols are size- and wavelength-dependent. The radiative effects were computed using Mie scattering for the standard Γ size distribution with effective variance of 0.2 for each size bin for reported^{8,9} refractive indices. Representative values of the radiative parameters at wavelength λ = 0.55 μm are given in Table 1. Dust radiative parameters are likely to vary regionally¹⁰, but the lack of detailed information about the global distribution of dust chemical composition and of dust refractive indices covering the full solar and thermal spectra prevent quantitative comparisons. Although nonsphericity of dust particles can produce large errors in remote-sensing applications based on Mie scattering computations, it has been shown by detailed T-matrix calculations that radiative fluxes and albedos utilizing Mie scattering results for equivalent volume spheres retain good accuracy¹¹.

Reflection, transmission and absorption of different specified aerosol distributions are calculated using the single gauss point

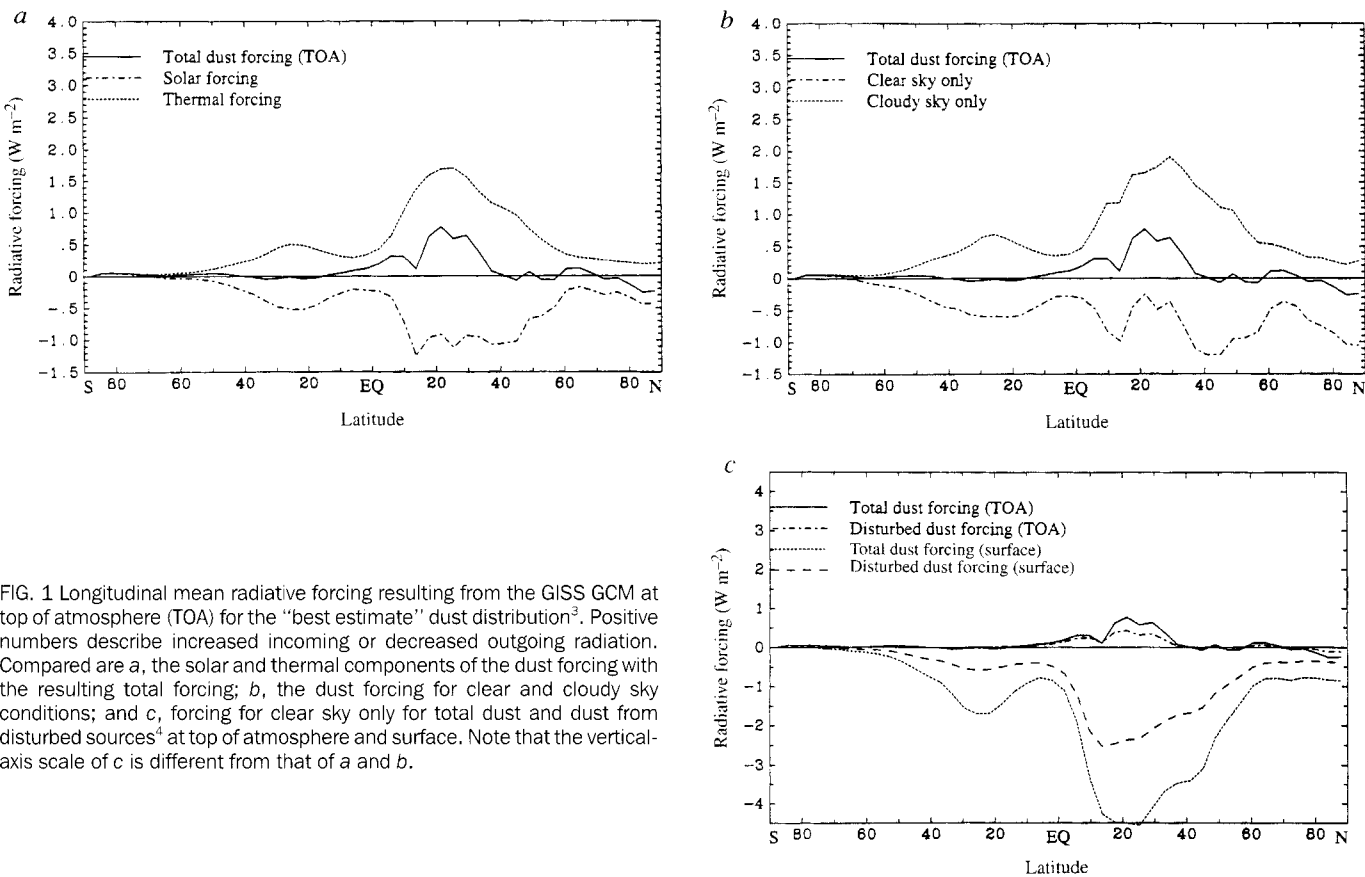


FIG. 1 Longitudinal mean radiative forcing resulting from the GISS GCM at top of atmosphere (TOA) for the “best estimate” dust distribution³. Positive numbers describe increased incoming or decreased outgoing radiation. Compared are a, the solar and thermal components of the dust forcing with the resulting total forcing; b, the dust forcing for clear and cloudy sky conditions; and c, forcing for clear sky only for total dust and dust from disturbed sources⁴ at top of atmosphere and surface. Note that the vertical-axis scale of c is different from that of a and b.

doubling/adding radiative transfer model in the GISS GCM^{12–14}. The correlated k-distribution method used in the GISS GCM to compute absorption by gases and particles is nominally accurate within 1% compared with line-by-line calculations¹⁵. The radiation model supports a continuous range of aerosol effective radii from 0.1 to 10 μm . The radiative forcing by dust was calculated using the modelled distributions for eight size classes.

The GCM was first run for one model year without dust aerosols in the atmosphere. It was then re-run for a model year with a specified aerosol distribution, with the atmospheric circulation and all other radiative constituents (for example, cloud optical properties and water vapour) prescribed by the no-dust experiment. The difference in the radiative fluxes between the aerosol and no-aerosol experiments defines the instantaneous radiative forcing with no dynamic feedback effects of the added dust load included.

At solar wavelengths, mineral aerosols are partly absorbing, with the single scattering albedo decreasing with increasing particle size^{14,16}. Whether or not the dust absorption results in an increase or decrease of the planetary albedo depends on the dust single scattering albedo as well as on the albedo of the underlying surface and on the cloud cover. Over dark surfaces like the ocean, the dust increases the planetary albedo because backscattering of incident solar radiation of the dust layer exceeds that from the dark surface. On the other hand, over high-albedo surfaces, such as

snow or bright deserts, the additional absorption by the dust reduces the solar flux reflected from the surface and results in a net reduction of radiation to space.

Figure 1a shows the solar and thermal components of the top-of-atmosphere (TOA) dust forcing. Solar forcing is mainly negative (cooling), as longitudinally the surface is darker than the dust. Maximum negative forcing occurs in areas of high dust optical thickness over the Arabian Sea, and to a lesser extent over the North Atlantic Ocean off the west coast of Africa. Thermal forcing is always positive as the dust absorption at thermal wavelengths contributes to greenhouse warming. The maxima of solar and thermal forcing occur at different latitudes because solar forcing over the bright Saharan and Arabian deserts and the dark ocean surfaces partially cancel around 20–30° N, where the dust optical thickness is highest. The total solar and thermal forcing is also nearly cancelled in the longitudinal mean.

TABLE 1 Effects of dust particle size

Effective radius (μm)	Q_{ext} (0.55 μm)	ω_0 (0.55 μm)	g (0.55 μm)	Fraction of TOA solar forcing (%)	Fraction of TOA thermal forcing (%)
0.1	0.644	0.964	0.515	2.4	0.3
0.2	2.316	0.973	0.663	27.1	2.0
0.4	3.086	0.956	0.666	52.5	6.0
0.8	2.583	0.907	0.688	28.8	18.7
1	2.476	0.887	0.716	2.3	68.4
2	2.280	0.816	0.797	–10.9	4.7
4	2.174	0.724	0.857	–4.2	1.0
8	2.110	0.633	0.904	–1.1	0.3

Radiation parameters and modelled contributions to the global solar and thermal TOA forcings for individual particle sizes. For the individual particle sizes extinction efficiencies Q_{ext} , single scattering albedos ω_0 , and symmetry parameters g are given for the solar wavelength 0.55 μm . A refractive index of $1.56 + 0.006i$ at this wavelength was used⁹. Uncertainties in the refractive index are²⁰ of the order of 40%. The percentage of contribution of individual particle sizes to solar and thermal dust forcing is given for total dust.

Table 1 summarizes the global mean fraction of solar and thermal forcing contributed by the different particle sizes. The TOA solar forcing becomes increasingly positive for particles larger than $1\text{ }\mu\text{m}$, because the larger particles absorb more strongly at solar wavelengths^{14,16}. The outgoing thermal radiation is dominated by particles with an effective radius of $1\text{ }\mu\text{m}$ as the smaller particles have smaller extinction coefficients and larger particles have short atmospheric lifetimes.

Figure 1b separates the total TOA dust forcing into clear-sky and cloudy-sky contributions. In the longitudinal mean, the clear-sky forcing is essentially negative because the aerosol is more reflecting than the ground surface, whereas the cloudy-sky forcing is positive owing to increased absorption of solar radiation by dust within and above the clouds. The maximum total forcing for clear skies is located at 40°N , even though the maximum dust loading is at $10\text{--}30^\circ\text{N}$. Again, this is due to the partial cancellation between total forcing over bright desert surfaces and the forcing over the darker oceans.

In Figure 1c we compare TOA and surface forcing for total dust and dust from disturbed sources only. Surface forcing is considerably larger than TOA forcing. Dust forcing at TOA is small in the longitudinal mean because of cancellation of the clear- and cloudy-sky contributions to the solar forcing. The TOA dust effect would be underestimated if only longitudinal means were considered.

Given the inherent uncertainties in dust refractive indices and particle sizes, the mean TOA forcing by mineral dust could easily be zero (or even slightly negative) if the aerosol particles were somewhat smaller in size or less absorbing. On the other hand, the substantial flux reduction at the surface and the increase in atmospheric absorption are the more robust characteristics of the effect of dust on climate.

To put the radiative forcing by disturbance dust in context with the forcings by greenhouse gases and industrial sulphate aerosols, we use the greenhouse forcing due to increases in CO_2 , CH_4 , N_2O and chlorofluorocarbons (CFCs) during the industrial era as calculated in the GISS GCM¹⁷. The global and annual mean forcing due to the increase in anthropogenic greenhouse gases that occurred from 1850 to 1980 is about $+2.1\text{ W m}^{-2}$ with an uncertainty of $\sim 10\%$.

Estimates of sulphate forcing are based on the regional pattern of anthropogenic sulphate distribution¹⁸ and radiative parameters¹⁷. The forcing can vary by more than a factor of 2 depending on different assumptions regarding the global sulphate burden and seasonal variations¹⁹. To facilitate comparison with other forcings and other studies, we scaled the resulting annual and global mean instantaneous sulphate forcing derived by the GISS GCM to a value of -0.28 W m^{-2} , which is equal to the value given in ref. 20.

In contrast to greenhouse gases, which have a positive thermal forcing effect, and sulphate aerosols, which cause negative solar forcing, dust forcing at solar wavelengths can produce either positive or negative forcing depending on clear/cloudy sky conditions and on surface albedo. Dust forcing at thermal wavelengths is always positive. The resultant total TOA radiative forcing by disturbance dust ranges locally from -2.1 to $+5.5\text{ W m}^{-2}$. The global mean net forcing at TOA calculated for wind-blown mineral dust is relatively small, with a mean value of $+0.14\text{ W m}^{-2}$ for total dust and $+0.09\text{ W m}^{-2}$ for dust from disturbed sources (which is a result of -0.25 W m^{-2} solar and $+0.34\text{ W m}^{-2}$ thermal forcing). This can be compared to the combined anthropogenic greenhouse gases and sulphate forcing of $+1.8\text{ W m}^{-2}$ in the global mean²⁰. The fact that the solar and thermal radiation forcings at TOA nearly cancel strongly suggests that mean radiative forcing at TOA does not provide sufficient information to fully characterize the effect of aerosol on climate.

Radiative forcing at the surface by dust from disturbed sources (Fig. 2a) is negative, as both the absorption and reflection of solar radiation by the dust reduce the solar radiation incident at the ground. The global mean forcing at the ground due to dust from

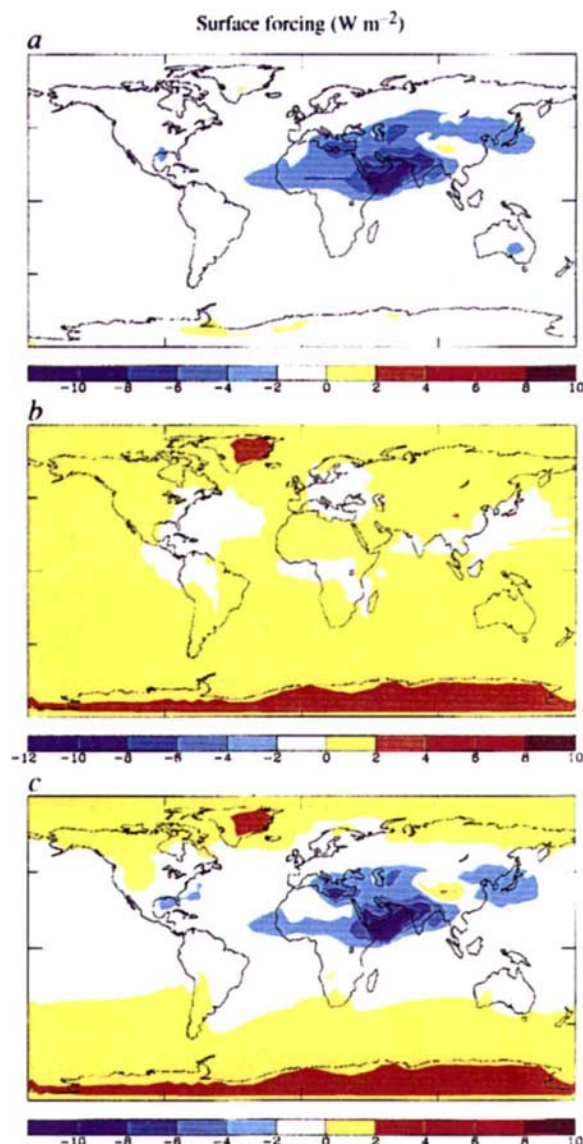


FIG. 2 Map of surface radiative forcing of dust from disturbed soils (a) compared with forcing caused by anthropogenic increase in greenhouse gases (CO_2 , CH_4 , N_2O and CFCs) and sulphate aerosols (b). For the combined forcing (c), the dust forcing was added to the anthropogenic greenhouse and sulphate forcing.

disturbed sources is -0.96 W m^{-2} , with a maximum of -25 W m^{-2} occurring over the Arabian Sea. If the contributions from anthropogenic greenhouse gas and sulphate forcings (Fig. 2b) are added, the resulting global mean flux at the surface is reduced to -0.6 W m^{-2} (Fig. 2c) compared to $+0.4\text{ W m}^{-2}$ for the case of only anthropogenic greenhouse gases and sulphate. Thus dust from disturbed sources actually changes the sign of the global mean anthropogenic forcing at the surface. It is not possible to translate the uncertainties of the model input parameters (global mean source strength, percentage dust from disturbed soils and refractive indices) directly into uncertainties of the radiative forcing because the global distribution of the radiation input parameters (refractive indices and sub-micrometre dust size distribution) is not well known and thus a meaningful range of uncertainty cannot be given based on the few existing measurements.

Changes in the atmospheric dust load also changes the atmospheric heating rates. An increase in longitudinal mean heating rate of 0.04 K d^{-1} occurs between 10° and 30°N corresponding to the location of maximum dust load. The maximum heating caused by

disturbed dust is $\sim 2 \text{ K d}^{-1}$, which is in good agreement with estimates (in refs 3 and 21) of increased heating rates between 1.6 and 15.6 K d^{-1} for a dust optical thickness of 1.5 at $0.55 \mu\text{m}$ wavelength. This additional atmospheric heating may influence atmospheric stability in regions with high dust loading, and lead to modifications of atmospheric circulation patterns and regional climate.

To improve estimates of the radiative forcing and to assess uncertainties, additional laboratory, field and satellite measurements are required. The laboratory measurements should include complex refractive indices for (especially) submicrometre-sized dust from different source regions. The refractive indices must cover all wavelengths of the full solar and thermal spectra. The satellite observations should make use of spectrally resolved upwelling radiation to distinguish amongst aerosol types and to retrieve optical depths.

To assess fully the magnitude and the climate implications of disturbed soil mineral dust requires a fully interactive GCM where the dust heating effects are allowed to influence atmospheric dynamics. Furthermore, as both the sources and the removal of dust are directly dependent on local climate, the changes in radiative forcing by dust may induce changes in the hydrological cycle and in regional circulation patterns that lead, in turn, to further changes in the dust sources and sinks. Thus there is a strong potential for highly interactive feedback effects. $\square$

Received 5 September 1995; accepted 21 February 1996.

- Houghton, J. T. et al. (eds) *Climate Change 1994, Radiative Forcing of Climate Change and an Evaluation of the IPCC IS92 Emission Scenarios* (Cambridge Univ. Press, 1994).
- Mitchell, J. F. B., Johns, T. C., Gregory, J. M. & Tett, S. F. B. *Nature* **376**, 501–504 (1995).
- Carlson, T. N. & Benjamin, S. G. *J. Atmos. Sci.* **37**, 193–213 (1980).
- Tegen, I. & Fung, I. *J. Geophys. Res.* **100**, 18707–18726 (1995).
- Tegen, I. & Fung, I. *J. Geophys. Res.* **99**, 22897–22914 (1994).
- Prather, M. J., McElroy, B., Wofsy, S. C., Russell, G. R. & Rind, D. *J. Geophys. Res.* **92**, 6579–6613 (1987).
- Fung, I., Prentice, K., Matthews, E. & Russell, G. *J. Geophys. Res.* **88**, 1281–1294 (1983).
- Volz, F. E. *Appl. Opt.* **12**, 564–568 (1973).
- Patterson, E. M., Gillette, D. A. & Stockton, B. H. *J. Geophys. Res.* **82**, 3153–3160 (1977).
- Sokolik, I., Andronova, A. & Johnson, T. C. *Atmos. Environ.* **27A**, 2495–2502 (1993).
- Mishchenko, M. I. *Appl. Opt.* **32**, 4652–4666 (1993).
- Hansen, J. E. & Travis, L. D. *Space Sci. Rev.* **16**, 527–610 (1974).
- Hansen, J. E. et al. *Mon. Weath. Rev.* **111**, 609–662 (1983).
- Lacis, A. A. & Mishchenko, M. I. In *Aerosol Forcing of Climate* (eds Charlson, R. & Heintzenberg, J.) 11–42 (Wiley, New York, 1995).
- Lacis, A. A. & Oinas, V. *J. Geophys. Res.* **96**, 9027–9064 (1991).
- Tegen, I. & Lacis, A. A. in *Proc. Int. Speciality Conf.* 293–305 (Air & Waste Mgmt Assoc., Pittsburgh, 1994).
- Hansen, J. E. et al. *Geophys. Res.* **93**, 9341–9364 (1988).
- Langner, J. & Rodhe, H. *J. Atmos. Chem.* **13**, 225–263 (1991).
- Kiehl, J. T. & Rodhe, H. in *Aerosol Forcing of Climate* (eds Charlson, R. & Heintzenberg, J.) 281–296 (Wiley, New York, 1995).
- Kiehl, J. T. & Briegleb, B. P. *Science* **260**, 311–314 (1993).
- Sokolik, I. & Golitsyn, G. *Atmos. Environ.* **27A**, 2509–2517 (1993).
- Li, X., Maring, H., Saville, D., Voss, K. & Prospero, J. M. *Nature* **380**, 416–419 (1996).

ACKNOWLEDGEMENTS. This work was supported by the NASA Earth Observing Systems (EOS) programme and by the US National Science Foundation. It is a contribution to the Climate System History and Dynamics Programme that is jointly sponsored by the Natural Sciences and Engineering Research Council of Canada and the Atmospheric Environment Service of Canada.

Reorganization of deep ocean circulation accompanying a Late Cretaceous extinction event

Kenneth G. MacLeod & Brian T. Huber

Department of Paleobiology, National Museum of Natural History, Smithsonian Institution, MRC: NHB 121, Washington DC 20560, USA

DEEP ocean circulation may be a significant factor in determining global climate^{1–5}. Increases in the flux of warm, saline waters from low to high latitudes would enhance the poleward transport of heat and, thus, help maintain the warm conditions at high latitudes typical of globally warm 'greenhouse' periods. But controversy exists^{1,2} as to whether the ocean's thermohaline circulation can transport enough heat to bring about the temperature distributions of these times, such as the mid-Cretaceous and early Eocene. Here we present stable-isotope records of ocean temperature and salinity that indicate that bottom waters in Late Cretaceous oceans of the Southern Hemisphere became cooler and less saline at the same time (about 70 Myr ago) as widespread biotic changes^{3,6–8}. These findings support the idea that changes in deep ocean circulation can act as a climate switch.

The Maastrichtian Age (the last 6 Myr of the Cretaceous Period) was a time of global environmental and biotic change. Isotopic data^{9–11}, microfossil palaeobiogeography¹², and palaeobotanical evidence^{13,14} all indicate cooling at this time. Several studies have shown warming during the last ~ 0.5 Myr of the Maastrichtian^{12,15,16} but superimposed on, not contradicting, evidence for long-term cooling. The disappearance of epicontinental seas as well as sequence stratigraphy¹⁷ document a large-scale, mid-Maastrichtian regression. Mid-Maastrichtian bioevents include floral shifts in the American west⁶, changes among benthic foraminiferal assemblages³, and extinction among both tropical, reef-forming rudistid⁷ and widely distributed, deep-sea inoceramid⁸ bivalves.

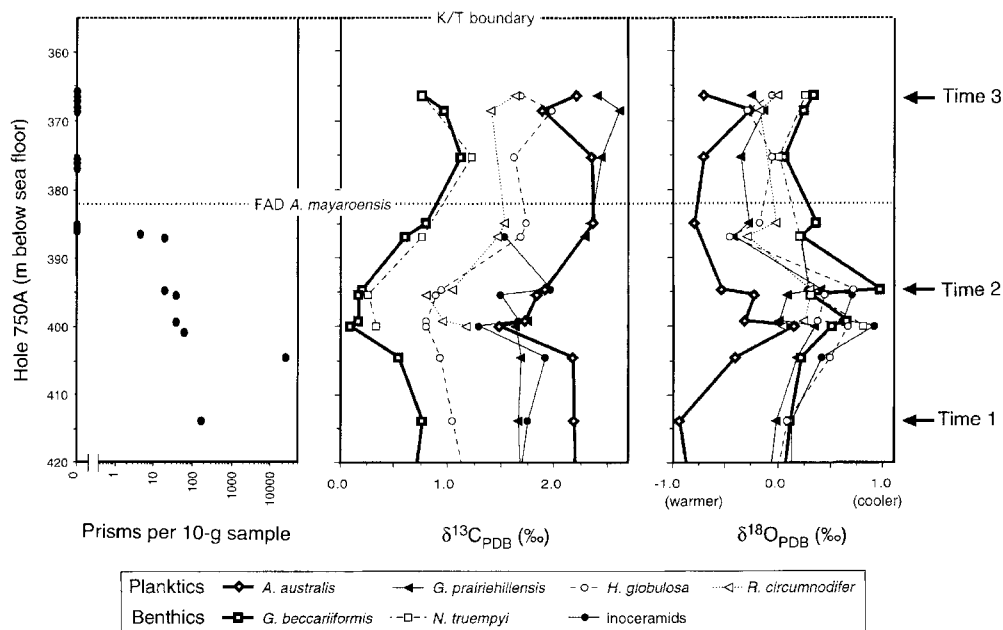
Oceanic processes may provide a causal mechanism uniting these events⁵. Global cooling and draining of tropical and subtropical epicontinental seas are expected to have reduced the production of hypersaline waters. Extinction among tropical, reef-

forming rudists has been attributed to the disappearance of hypersaline surface waters⁷; habitat restriction related to decreased influence of saline bottom waters has been invoked to explain extinction among deep-sea inoceramids⁸. Because it would have decreased oceanic heat transport, declining importance of warm, saline water masses could have affected Maastrichtian climates and indirectly led to observed changes among terrestrial floras. But direct data supporting the hypothesized palaeoceanographic changes are lacking.

To constrain better the physical properties of Maastrichtian oceans, we measured $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ of size-sorted, single-taxon separates of planktic foraminifers, benthic foraminifers, and inoceramid remains from an austral (Hole 750A, $\sim 60^\circ\text{S}$ palaeolatitude) and a subtropical (Hole 761, $\sim 35^\circ\text{S}$ palaeolatitude) Ocean Drilling Program site. Preservation is uniformly good to excellent throughout the intervals studied, and local evidence of biological change (the disappearance of inoceramids) is well documented at both localities⁸. Because the inoceramid extinction is recognized globally⁸, isotopic shifts consistently associated with this event can be justifiably proposed as global phenomena. Similar-sized foraminifers were selected for each taxon analysed minimizing the possibility of size-related artefacts compromising the results¹⁸. Measured isotopic values are not significantly correlated with small variations in size that existed among samples of any taxon.

At both sites there is a $\sim 0.7\text{‰}$ negative $\delta^{13}\text{C}$ excursion in benthic foraminifers associated with the extinction of inoceramids (Figs 1 and 2). The abundance of inoceramid shell fragments (prisms) begins to decline at the initiation of this excursion, and inoceramid disappearance coincides with the point where $\delta^{13}\text{C}$ values return to original levels. Over the same stratigraphic interval there is a $0.5\text{--}1.0\text{‰}$ positive shift in $\delta^{18}\text{O}$ values. Change at the subtropical site is more or less unidirectional; at the austral site, values first become heavier and then largely recover. Results from inoceramid prisms parallel the trends shown by benthic foraminifers, but exhibit consistent offsets perhaps related to chemosymbiosis in inoceramids or vital effects in benthic foraminifers^{19–21}. Incorporating planktic data in the analysis shows that correlative changes in surface waters occurred only at the austral site. Isotopic trends in austral planktic and benthic foraminifers are parallel (Fig. 1), whereas there is no obvious correlation between planktic and benthic results in either $\delta^{13}\text{C}$ or $\delta^{18}\text{O}$ in the subtropical samples (Fig. 2).

FIG. 1 Inoceramid abundance and stable-isotope analyses $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ through the Maastrichtian in ODP Hole 750A ($\sim 60^\circ\text{S}$ palaeolatitude). Data for a benthic taxon and a shallow planktic taxon (based on average oxygen values) are highlighted. All taxa exhibit similar trends with a negative $\delta^{13}\text{C}$ and a positive $\delta^{18}\text{O}$ excursion associated with the decline of inoceramids. The seemingly anomalous inoceramid point on the $\delta^{13}\text{C}$ plot (the highest inoceramid point) could not be re-analysed because all available prisms were used in the first analysis and the next-higher sample did not contain a sufficient mass of prisms for an analysis. The dotted line labelled "FAD *A. mayaroensis*" is the first appearance datum of *Abathomphalus mayaroensis*, a planktic foraminifer commonly used as a zonal indicator for the upper Maastrichtian. Time slices 1, 2, and 3 are modelled in Fig. 3. Complementary data from this hole are presented in ref. 22. ($\delta^{13}\text{C} = [(^{13}\text{C}/^{12}\text{C})_{\text{sample}} / (^{13}\text{C}/^{12}\text{C})_{\text{standard}}] - 1$ and $\delta^{18}\text{O} = [(^{18}\text{O}/^{16}\text{O})_{\text{sample}} / (^{18}\text{O}/^{16}\text{O})_{\text{standard}}] - 1$; the PDB standard is used in both cases).



Carbon isotope data demonstrate mid-Maastrichtian palaeoceanographic change associated with inoceramid extinction but fail to yield a unique interpretation regarding the nature of that change. Further, competing hypotheses seem to make restrictive predictions. An explanation invoking a single cause for observed trends in $\delta^{13}\text{C}$ suggests a simple physical oceanographic relationship between the two sites. On the other hand, invoking separate benthic and planktic processes to explain the results requires that independent causes yielded $\delta^{13}\text{C}$ excursions of similar magnitude and direction. The lack of planktic $\delta^{13}\text{C}$ trends in subtropical samples (Fig. 2) argues against Maastrichtian redistribution of carbon among global reservoirs²²; therefore, planktic shifts at the austral site (Fig. 1) probably represent local changes in produc-

tivity. Sinking of surface-water carbon could explain parallel trends in co-occurring benthic taxa at this site. The $\delta^{13}\text{C}$ record of subtropical benthics, though, is clearly not tied to local surface values and requires either horizontal import of isotopic changes or temporal variation in benthic cycling of carbon. If the subtropical benthic $\delta^{13}\text{C}$ excursion reflects characteristics inherited from the source region(s) for bottom water, similarities between the two sites imply that they shared a single bottom-water mass (at any given time) with isotopic continuity over significant distances—differences in benthic $\delta^{13}\text{C}$ values are too small to infer transport direction^{23,24}. Alternatively, the subtropical (and austral) benthic record might indicate that changes in bottom waters (for example, temporal increase in the concentration of dissolved oxygen) led to

FIG. 2 ODP Hole 761 ($\sim 35^\circ\text{S}$ palaeolatitude). Benthic taxa exhibit the same isotope patterns associated with the inoceramid extinction observed in Hole 750A (Fig. 1), but the planktic taxa show little change across the interval studied. See text for further discussion. Asterisk indicates data from ref. 22.

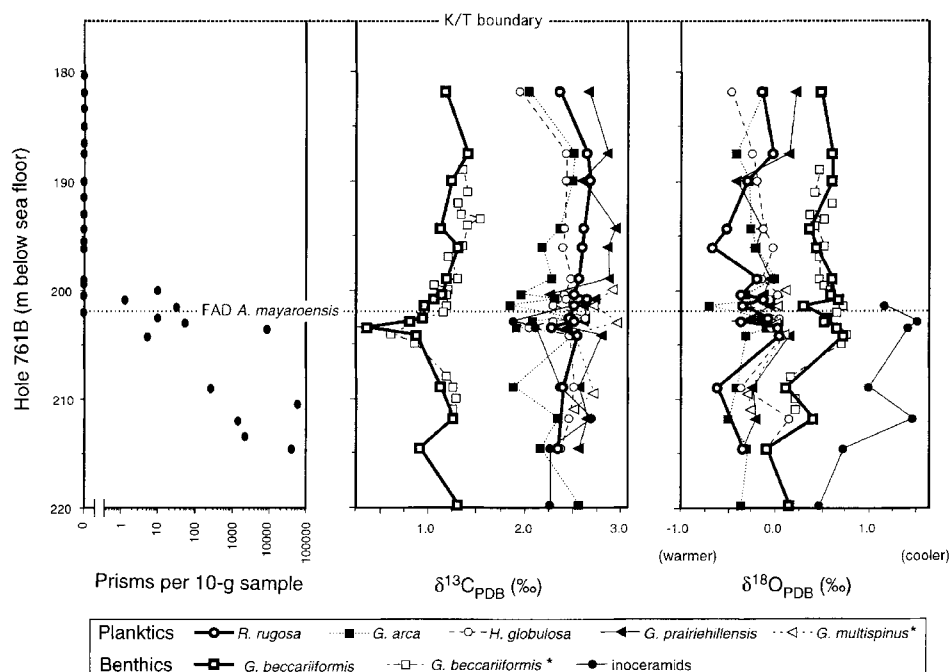
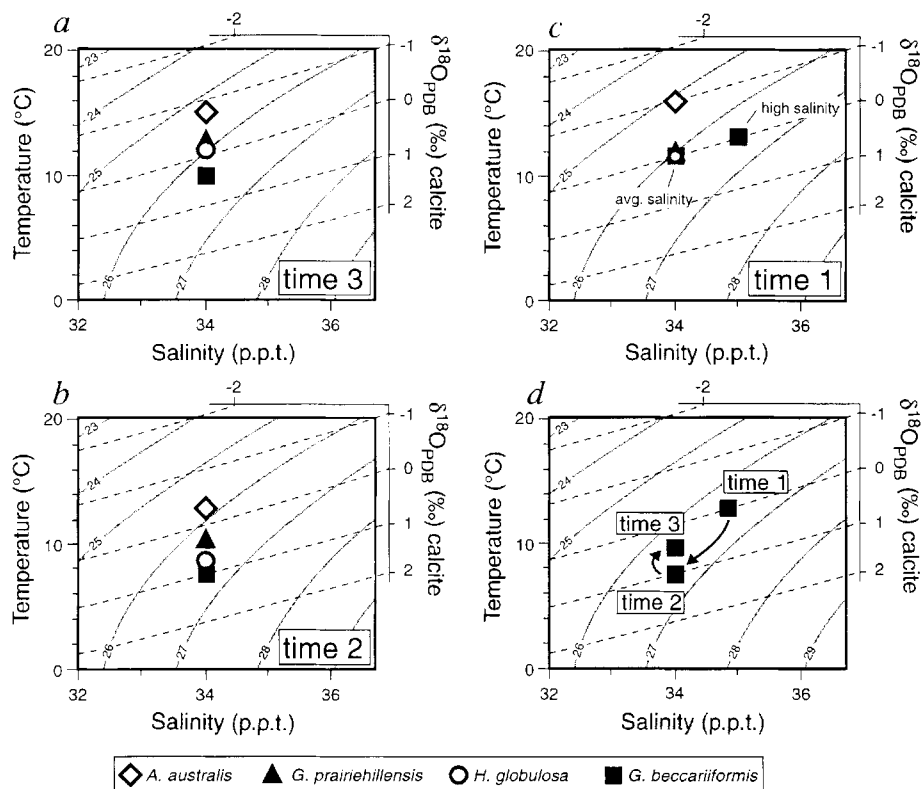


FIG. 3 $\delta^{18}\text{O}$ values of selected taxa from time slices 1, 2, and 3 (as indicated on Fig. 1) plotted in the temperature–salinity space appropriate for the Cretaceous period (after ref. 26). Sloping dashed lines represent $\delta^{18}\text{O}$ values of calcite in equilibrium with sea water of a given temperature and salinity. Curved solid lines are isopycnals expressed as σ_t (that is, $1.000(\rho_{\text{seawater}} - 1)$, where ρ is density in g cm^{-3}). a, Plot of time slice 3 (post-extinction) from Fig. 1, assuming an isohaline water column (34 p.p.t. is used as an appropriate Cretaceous value after ref. 26) showing apparent decreasing temperature and increasing density with depth. b, Plot of time slice 2 (inoceramids declining) suggesting cooler temperatures than time slice 3 but a stable, isohaline water column. c, Plot of time slice 1 (pre-extinction) suggesting a meta-stable water column if the water column were isohaline (*Gavellinella beccariiformis* symbol labelled “avg. salinity”) and a stable water column if time slice 1 bottom water had elevated salinity and temperatures (*G. beccariiformis* symbol labelled “high salinity”). d, Proposed changes in the relative temperature and salinity of bottom waters based on the analyses of *G. beccariiformis* from time slices 1, 2 and 3. See text for further discussion.



remineralization of previously deposited, isotopically light organic carbon perturbing the vertical gradient in $\delta^{13}\text{C}$. This interpretation is less restrictive in terms of physical oceanography, but it invokes two causes for one pattern. The similarities in shape, but not timing, between the $\delta^{13}\text{C}$ curve for austral planktics and one or both of the benthic records would be coincidental.

Oxygen isotope results provide better constraints on the nature of oceanographic change. Although typically used to estimate palaeotemperature, $\delta^{18}\text{O}$ analyses can also indicate changes in salinity. Evaporation preferentially removes water containing ^{16}O , so saline waters generated by evaporative concentration are enriched in ^{18}O . Taxa living in high salinity, ^{18}O -enriched waters secrete an isotopically heavier test than those living in waters of lower salinity but the same temperature. If both the isotopic composition of the oceans and the isotopic enrichment related to evaporation and freshwater input are known, the $\delta^{18}\text{O}$ values of calcite in equilibrium with these waters can be plotted in temperature/salinity space^{25,26}. Further, because water density varies as a function of temperature and salinity, the plots can be contoured for density (isopycnals). By estimating the unknown variables, Woo *et al.* (ref. 26) constructed such a plot appropriate for Late Cretaceous oceans. Errors in estimated parameters would shift the position of the equilibrium lines, but would not greatly affect inferred differences in temperature/salinity among samples.

The $\delta^{18}\text{O}$ analyses suggest that bottom-water source regions switched from low to high latitudes coincident with the inoceramid extinction at both sites. At the austral site, if the water column was isohaline, $\delta^{18}\text{O}$ results for time 2 and time 3 (see Fig. 1) indicate a stable water column in which temperature decreases and density increases with depth (Fig. 3). However, at time 1, if the water column was isohaline, temperature would not change below the surface layer and the water column would be only metastable; further, some analyses between times 1 and 2 indicate density inversion. This paradox disappears if the pre-extinction bottom waters had higher salinity and temperature. Assuming bottom-water density at time 3 was equal to bottom-water density at time 1 (that is, cool, post-extinction bottom waters were as dense as the warm, salty bottom waters they are proposed to have replaced),

the $\delta^{18}\text{O}$ value of *Gavellinella beccariiformis* at time 1 suggests a stable water column in which pre-extinction bottom waters were ~ 1 part per thousand (p.p.t.) saltier and $\sim 1.5^\circ\text{C}$ warmer than deep surface waters. The subtropical site gives similar results, but lacks the evidence for a peak in cooling analogous to time 2 at the austral site. Thus, initiation or intensification of intermediate and/or deep-water formation in the austral region may coincide with the cooling at time 2. These new, cool, dense waters then progressively displaced warm, saline waters at depth without demonstrably affecting mid-latitude surface water.

The salinity changes proposed above are conservative. If the density of bottom water at time 1 is estimated from results for time 2 rather than time 3, modelled pre-extinction bottom water is an additional ~ 0.5 p.p.t. saltier and $\sim 1^\circ\text{C}$ warmer. The assumption of density balance across time, though, is not required. For example, if the production of hypersaline waters declined or ceased in low-latitude seas, the extent of deep-sea regions bathed in warm, saline waters would have contracted towards the Equator without high-latitude forcing (implicitly assumed above). Thus, the upper limit of the salinity/temperature of pre-extinction bottom waters is bounded only by reasonable estimates of the ecological tolerances of contemporary benthic organisms. In addition, the possibility of a contribution from low-latitude forcing in reorganization of deep-ocean circulation means observed austral temperature changes may be more an effect than a cause of Maastrichtian change.

Uncertainty regarding cause and effect notwithstanding, we think a change in the location of the dominant source regions of bottom waters could have provided a causal mechanism that temporally focused Maastrichtian biotic changes. Whereas Late Cretaceous cooling occurred over an extended interval^{9,11,27}, palaeontological and geochemical shifts are concentrated in the mid-Maastrichtian^{7,8,22,28}. If the palaeoceanographic model presented above is accurate (that is, if a reversal in deep ocean circulation is coincident with global palaeoecological change) change in deep ocean circulation may have acted as a climate switch. A shift from low-latitude to high-latitude production of bottom waters is expected to have caused a decline in oceanic heat transport and lead to an increase in latitudinal temperature

gradients^{1,2}. Late Maastrichtian temperature gradients were steep relative to earlier in the Late Cretaceous¹¹, but additional data are needed to test whether this increase in gradient was correlated with mid-Maastrichtian changes. These constraints on the role and behaviour of the Maastrichtian deep ocean may have relevance to modern climate change. The reorganization of ocean circulation proposed here occurred when estimated atmospheric concentrations of CO₂ were about twice the present-day levels³⁹; the same CO₂ concentrations are employed in model simulations attempting to predict the next century's climate³⁰. Documenting and, more importantly, understanding controls on changes in ancient climates under these conditions could be crucial to making accurate predictions. □

Received 28 September 1995; accepted 16 February 1996.

1. Barron, E. J., Fawcett, P. J., Peterson, W. H., Pollard, D. & Thompson, S. L. *Paleoceanography* **10**, 953–962 (1995).
2. Sloan, L. C., Walker, J. C. G. & Moore, T. C. Jr *Paleoceanography* **10**, 347–356 (1995).
3. Thomas, E. in *Proc. ODP Sci. Res.* **113**, 571–594 (1990).
4. Kennett, J. P. & Stott, L. D. *Nature* **353**, 225–229 (1991).
5. Thomas, E. & Shackleton, N. J. in *Correlations of the Early Paleogene in Northwestern Europe* (eds Knox, R. O. B., Corfield, R. & Dunnay) (Geological Soc. London, in the press).
6. Johnson, K. R. & Hickey, L. J. (abstr.) *Meeting on Mesozoic of the Western Interior* 36 (Soc. of Econ. Paleontologists & Mineralogists, Ft Collins, Colorado, 1992).
7. Johnson, C. C. & Kauffman, E. G. in *Extinction Events in Earth History* (eds Kauffman, E. G. & Walliser, O. H.) 305–324 (Springer, Berlin, 1990).
8. MacLeod, K. G., Huber, B. T. & Ward, P. D. in *The Cretaceous-Tertiary Event and Other*

- Catastrophes in Earth History* (eds Ryder, G., Gartner, S. & Fastovsky, D.) (Spec. Pap., Geological Soc. Am., Boulder, in the press).
9. Barrera, E., Huber, B. T., Savin, S. M. & Webb, P. N. *Paleoceanography* **2**, 21–47 (1987).
 10. Barrera, E. & Huber, B. T. in *Proc. ODP Sci. Res.* **113**, 813–823 (1990).
 11. Huber, B. T., Hodell, D. A. & Hamilton, C. P. *Geol. Soc. Am. Bull.* **107**, 1164–1191 (1995).
 12. Huber, B. T. & Watkins, D. K. in *The Antarctic Paleoenvironment: A Perspective on Global Change* Pt 1 (eds Kennett, J. P. & Wamke, D. A.) Vol. 56, 31–60 (Antarctic Res. Ser., American Geophysical Union, Washington, D.C., 1992).
 13. Wolfe, J. A. & Upchurch, F. R. Jr *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **61**, 33–77 (1987).
 14. Spicer, R. A. & Parrish, J. T. *Palaeontology* **33**, 225–242 (1990).
 15. Johnson, K. R. & Hickey, L. J. in *Global Catastrophes in Earth History* (eds Sharpton, V. L. & Ward, P. D.) 433–444 (Spec. Pap. 247, Geological Soc. Am., Boulder, 1990).
 16. Stott, L. D. & Kennett, J. P. in *Proc. ODP Sci. Res.* **113**, 829–848 (1990).
 17. Haq, B. V., Hardenbol, J. & Vail, P. R. *Science* **235**, 1156–1167 (1987).
 18. Berger, W. H., Killingley, J. S. & Vincent, E. *Oceanologica Acta* **1**, 203–216 (1978).
 19. MacLeod, K. G. & Hoppe, K. A. *Geology* **20**, 117–120 (1992).
 20. MacLeod, K. G. & Hoppe, K. A. *Geology* **21**, 95–96 (1993).
 21. Grossman, E. L. *Geology* **21**, 94–95 (1993).
 22. Barrera, E. *Geology* **22**, 877–880 (1994).
 23. Curry, W. & Lohmann, G. P. *Quat. Res.* **18**, 218–235 (1982).
 24. Shackleton, N. J., Imbrie, J. & Hall, M. A. *Earth planet. Sci. Lett.* **65**, 233–244 (1983).
 25. Railsback, L. B., Anderson, T. F., Ackerly, S. C. & Cisne, J. L. *Paleoceanography* **4**, 585–591 (1989).
 26. Woo, K. S., Anderson, T. F., Railsback, L. B. & Sandberg, P. A. *Paleoceanography* **7**, 673–685 (1992).
 27. Spicer, R. A. & Corfield, R. M. *Geol. Mag.* **129**, 169–180 (1992).
 28. Nelson, B. K., MacLeod, K. G. & Ward, P. D. *Nature* **351**, 644–647 (1991).
 29. Berner, R. A. *Am. J. Sci.* **294**, 56–91 (1994).
 30. Barron, E. J. *Eos* **76**, 185–191 (1995).

ACKNOWLEDGEMENTS. We thank A. Cutler, M. Ducharme and S. Wise for comments on early drafts of this Letter; the Ocean Drilling Program for providing samples; and D. Hodell for isotopic analyses performed at the University of Florida. This work was supported by the Walcott Fund and the Smithsonian Institution's Museum of Natural History.

Creep events preceding small to moderate earthquakes on the San Andreas fault

Clifford H. Thurber

Department of Geology and Geophysics, University of Wisconsin-Madison, 1215 West Dayton Street, Madison, Wisconsin 53706, USA

THE San Andreas fault in central California is unusual in that much of the fault displacement occurs as aseismic slip, also known as creep. Changes in creep have been suggested as a possible earthquake precursor^{1–6}, but clear associations between creep changes and subsequent earthquakes are uncommon. Here I report a test of the hypothesis that episodes of rapid creep ('creep events'^{1–3}) precede small to moderate earthquakes on the creeping portion of the San Andreas. The test consists of a comparison of archival creep data to the earthquake catalogue for this area, and an earthquake prediction experiment. Between 1970 and 1994, creep events occurred in the five days preceding half or more of the earthquakes of magnitude ≥ 3.3 in three 1-yr periods of above-average seismic activity. Of five predictions based on the occurrence of creep events between October 1995 and mid-January 1996, four were fulfilled; there were also two 'misses' (earthquakes not preceded by a creep event). If the relationship reported here holds for the entire creeping segment of the San Andreas, and applies to larger earthquakes, these results suggest that a prediction based on precursory creep may be possible at Parkfield, California.

Creep can involve relatively steady motion or more abrupt episodes called creep events^{1–3}. Creep events are thought to be restricted to the very shallow crust, probably no deeper than 1 or 2 km (refs 1–3), whereas the seismogenic zone in central California extends from about 2 to 10 km depth^{7,8}. Creep events and small to moderate earthquakes in this area appear to be recurring phenomena on specific patches of the shallow and deeper fault, respectively^{1–4,7,8}.

In April to September 1995, five successive small to moderate earthquakes (local magnitude $M_L = 3.7$ –5.0) along a 40-km stretch of the creeping portion of the San Andreas fault were preceded by creep events detected on creepmeters in the area operated by the US Geological Survey⁹ (USGS) (Figs 1 and 2, Table 1). The time interval between the creep events and the subsequent earthquakes ranged from 1 to 5 days, and the creep event amplitudes ranged from about 0.5 to 3.0 mm. There was no direct spatial association between the creep events and the earthquakes. Five other creep events (one a pair of creep events coincident in time) occurred during this time period (Fig. 2). Thus from April to September, there were five instances when creep events were followed within 5 days by small to moderate earthquakes (five 'hits'), and five instances when they were not (five 'false alarms'). In contrast, there were post-seismic creep events but no significant precursory creep associated with two small earthquakes in early 1995 (M_L 3.3, 7 January; M_L 3.5, 16 February) (Figs 1 and 2, Table 1).

These examples of creep events preceding small to moderate earthquakes stand in sharp contrast to the bulk of the creep literature. Associations between creep events and earthquakes have been reported in a number of studies, but virtually all cases involved creep events triggered by earthquakes, not precursory

TABLE 1 Source parameters for earthquakes in Fig. 1 and associated creep events

No.	Date (y/m/d), time (ut)	Lat. (°N)	Lon. (°W)	Depth (km)	Magnitude	Comment; observing creepmeter
1	97/01/07, 14:53:40	36.81	121.54	7.2	3.3	Post-seismic creep; XMR
2	95/02/16, 00:36:22	36.80	121.52	6.3	3.5	Post-seismic creep; XSJ
3	95/04/23, 08:41:36	36.60	121.20	7.3	5.0	Creep precursor; XHR, CWN
4	95/06/16, 04:20:36	36.75	121.37	9.9	3.9	Creep precursor; XMR
5	95/07/12, 21:58:00	36.59	121.20	8.9	4.0	Creep precursor; XSJ
6	95/08/29, 23:08:58	36.69	121.32	6.6	3.7	Creep precursor; XHR, XMR
7	95/09/27, 16:44:42	36.58	121.17	9.3	4.1	Creep precursor; CWN

The source for this earthquake data was the Northern

TABLE 2 Creep events and earthquakes from hypothesis-testing periods

Period (y/m/d), event type	Date (y/m/d), time (ut)	Location (Lat. °N, Lon. °W, or site)	Comment*
72/02/01–73/01/31			
Earthquake	72/02/18, 05:55	36.77, 121.49	Miss, M_L 3.4
Earthquake	72/02/22, 05:41	36.55, 121.11	Miss, M_L 3.5
Creep event	72/02/23	XMR	Precursory creep
Earthquake	72/02/24, 15:56	36.59, 121.20	Hit, M_L 5.1
Creep event	72/03/12	XMR	Precursory creep
Earthquake	72/03/13	36.67, 121.30	Hit, M_L 3.3
Earthquake	72/03/14	36.94, 121.69	Hit, M_L 3.3
Creep event	72/04/27	XMR	False alarm
Creep event	72/05/22	XSJ	False alarm
Creep event	72/07/01	XMR	False alarm
Creep event	72/09/02	XMR	Precursory creep
Earthquake	72/09/04, 18:04	36.64, 121.25	Hit, M_L 4.7
Earthquake	72/09/23, 15:07	36.80, 121.51	Miss, M_L 4.1
Creep event	72/10/02	XSJ	Precursory creep
Earthquake	72/10/03, 06:30	36.80, 121.52	Hit, M_L 4.8
Creep event	72/10/10	XMR	False alarm
Creep event	72/10/17	XSJ	False alarm
Creep event	72/10/25	XMR	False alarm
Creep event	73/01/10	XMR	Precursory creep
Earthquake	73/01/15, 09:43	36.68, 121.31	Hit, M_L 4.1
85/06/01–86/05/31			
Creep event	85/06/21	XHR	False alarm
Earthquake	85/06/27, 04:38	36.53, 121.09	Miss, M_L 3.4
Creep event	85/07/24	XMR	False alarm
Creep event	85/09/21	XFL	False alarm
Creep event	85/10/13	CWN, XHR	False alarm
Creep event	85/10/21	CWN	Precursory creep
Earthquake	85/10/26, 14:30	36.83, 121.57	Hit, M_L 3.4
Creep event	85/11/22	XMR	False alarm
Creep event	86/01/10	XHR	Precursory creep
Earthquake	86/01/14, 03:09	36.58, 121.18	Hit, M_L 4.8
Creep event	86/01/26	CWN	False alarm
Creep event	86/02/16	XMR	False alarm
Earthquake	86/04/15, 09:25	36.69, 121.32	Miss, M_L 3.6
Creep event	86/05/30	CWN	Precursory creep
Earthquake	86/05/31, 08:47	36.64, 121.25	Hit, M_L 4.7
90/08/01–91/07/31			
Creep event	90/08/01	XHR	Precursory creep
Earthquake	90/08/05, 06:52	36.88, 121.63	Hit, M_L 4.0
Earthquake	90/08/25, 17:39	36.68, 121.31	Miss, M_L 3.5
Creep event	90/08/31	XSJ	Precursory creep
Earthquake	90/09/02, 08:08	36.67, 121.29	Hit, M_L 3.5
Earthquake	90/09/10, 01:43	36.89, 121.63	Miss, M_L 3.5
Earthquake	90/09/11, 07:14	36.67, 121.29	Miss, M_L 4.4
Earthquake	90/09/23, 03:00	36.80, 121.53	Miss, M_L 3.7
Creep events	90/10/09, 90/10/12	CWN, XHR	False alarm
Creep events	90/10/19	CWN, XHR	False alarm
Creep event	90/11/02	XMR	False alarm
Creep event	90/11/23	XSJ	False alarm
Creep event	91/02/04	XMR	False alarm
Creep event	91/03/10	XFL, XSJ	False alarm
Creep event	91/03/19	CWN	Precursory creep
Earthquake	91/03/24, 03:42	36.96, 121.73	Hit, M_L 4.6
Creep event	91/06/06	CWN	Precursory creep
Earthquake	91/06/08, 13:40	36.85, 121.58	Hit, M_L 4.1
Creep event	91/07/22	XFL	False alarm
95/10/01–96/01/15			
Creep event	95/11/09, 11:30	XMR	Precursory creep
Earthquake	95/11/12, 20:58	36.68, 121.36	Hit, M_L 3.8
Creep event	95/11/19, 08:30	CWN	False alarm
Earthquake	95/11/25, 00:43	36.70, 121.33	Miss, M_L 3.3
Creep event	95/11/27, 11:30	CWN	Precursory creep
Earthquake	95/11/29, 23:12	36.70, 121.33	Hit, M_L 4.2
Earthquake	95/12/25, 06:32	36.81, 121.55	Miss, M_L 3.3
Creep event	96/01/07, 23:00	XSJ	Precursory creep
Earthquake	96/11/10, 02:16	36.80, 121.54	Hit, M_L 3.6
Creep event	96/01/14, 20:30	XMR	Precursory creep
Earthquake	96/01/15, 01:41	36.69, 121.32	Hit, M_L 3.4

* See text for definition of 'hit', 'miss' and 'false alarm'.

changes in creep. Some exceptions include a report of creep retardation before earthquakes on this section of the San Andreas Fault from 1971 to 1987², evidence for accelerated fault slip before the 1966 Parkfield earthquake^{10,11}, and a report of premonitory slip preceding the Landers earthquake¹². There are several possible explanations for this discrepancy. One is that most previous studies may have used a restrictive spatial window for seeking associations between creep events and earthquakes, and thus failed to identify precursory creep. Note that for the associations identified in Fig. 2, the creepmeter nearest to each earthquake does not show precursory creep. Another possibility is that some time-dependent process has led to a change in the relationship between creep events and seismicity. The year 1995 was unusual in that the level of seismic activity was significantly above average, suggesting that creep events might be precursory at times when a significant portion of the fault segment is near failure, as indicated by a high level of seismic activity.

I accordingly tested the hypothesis that creep events occur as precursors to small to moderate earthquakes in periods of above-average seismic activity—by comparing the USGS archive of creep data to the earthquake catalogue for this area, and by carrying out an earthquake prediction experiment beginning 1 October 1995. The 1-yr time periods of (A) February 1972 to January 1973, (B) June 1985 to May 1986, and (C) August 1990 to July 1991 were selected as example time periods for analysis based on their heightened seismic activity levels (1-yr periods were used to allow for an adequate statistical test). The period immediately following the 1989 Loma Prieta earthquake also had above-average seismicity, but creep in the region was profoundly disturbed by that earthquake¹³, so that period was not considered. A relatively low magnitude threshold of 3.3 was chosen; in general, a higher threshold will decrease the miss rate but increase the false alarm rate¹⁴. During these three periods, a majority of the $M_L \geq 3.3$ earthquakes were preceded by creep events within five days (Table 2).

In period A, there were six hits, three misses and six false alarms; in period B there were three hits, two misses and six false alarms; in period C there were four hits, four misses and seven false alarms (Table 2). I note that although the earthquake catalogue is relatively complete down to $M_L \approx 1.0$ in this area, the sparseness of the creepmeter network suggests that an unknown number of creep events may go undetected. Thus it is not surprising that there are numerous misses. The probability of observing by chance K or more out of L earthquakes of $M_L \geq 3.3$ occurring in 5-day windows following N creep events, assuming that the earthquakes occur

randomly (Poisson behaviour), is given by

$$P = 1 - \sum_{i=1}^{K-1} E^i \exp(-E)/(i!)$$

where E is $(N \times 5 \text{ days}) \times (L/365 \text{ days})$. The probabilities are 0.003 for period A, 0.024 for period B and 0.034 for period C. These results indicate that the apparent temporal relation between creep events and subsequent small to moderate earthquakes is unlikely to be a coincidence.

For the earthquake prediction experiment, I issued a prediction each time a creep event occurred, indicating my expectation that an $M_L \geq 3.3$ earthquake would occur within 5 days of the creep event. Between 1 October 1995 and 15 January 1996, there were five creep events resulting in predictions (Table 2); four were hits and the other was a near miss (an $M_L = 3.3$ earthquake occurred 5 days and 16 hours after a creep event on 19 November; a false alarm and a miss). In addition, an $M_L = 3.3$ earthquake occurred on 25 December with no precursory creep event (a miss; Table 2). The probability that four or more hits (out of six earthquakes and five creep event "chances") in the 107-day period would occur randomly is very nearly significant at $P = 0.054$.

These observations potentially have a number of significant implications. In the short term, the occurrence of further creep events on this part of the San Andreas fault may permit additional successful predictions. However, the period of precursory creep events can be expected to end in the near future. In the longer term, a study of the relationship between creep events and seismicity in this area may yield important insights into the mechanical behaviour of the San Andreas fault. For example, successful fault models should be able to account for the variability of the temporal relationship between creep events and earthquakes. The creep events and earthquakes may be triggered in common by regional changes in the state of stress, in which case

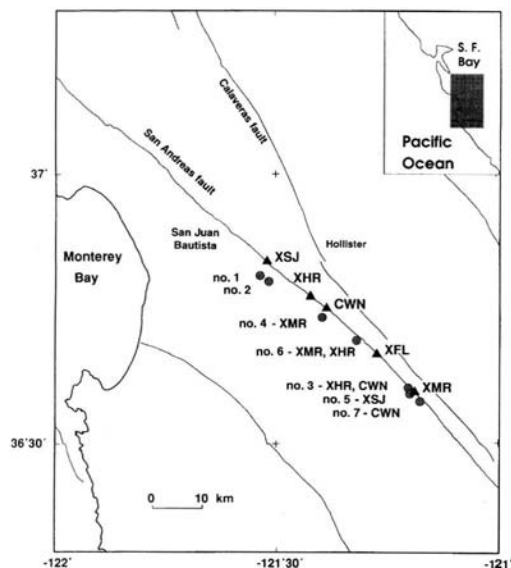


FIG. 1 Map showing locations of USGS creepmeters (triangles with three-letter station code names), data from which were analysed. Also shown are the locations of the January–September 1995 earthquakes along the San Andreas fault that are listed in Table 1 (filled circles) in the magnitude range 3.3–5.0 (excluding aftershocks), and their identification numbers, noting creepmeters with precursory creep (if any), and mapped faults (solid lines). Inset, map location in central California. The creepmeters consist of invar wires, rods or beams crossing the fault at an angle of 22–45°, anchored at one end, and attached to a transducer or potentiometer at the other end to measure displacement⁹. Creepmeter XFL was not operational during 1995.

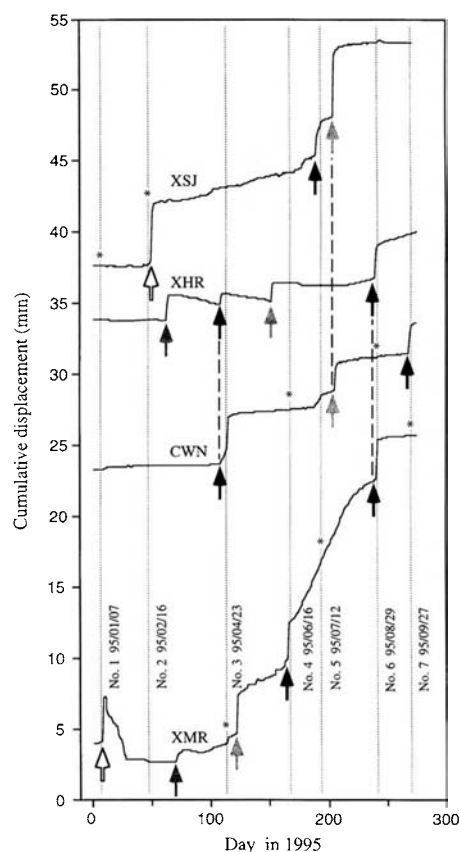


FIG. 2 Records of cumulative displacement for the four operative USGS creepmeters (curves marked with the three-letter station code names) and times of the magnitude 3.3–5.0 earthquakes (dotted vertical lines) shown in the map in Fig. 1 and listed in Table 1. An asterisk indicates the creep record for the creepmeter closest to the earthquake. The creep curves have been offset for clarity. Creep events were defined as one-day displacements that exceeded the main daily creep rate plus one standard deviation for a given creepmeter for the time period considered. A total of 15 creep events were identified (arrows)—dashed lines connect three pairs of coincident creep events. Beginning with 23 April earthquake, each small to moderate earthquake was preceded by a creep event at one or more creepmeters (seven black arrows). In contrast, the first two earthquakes in 1995 (numbers 1 and 2, magnitudes 3.3 and 3.5) showed post-seismic creep events (two unfilled arrows) but no significant precursory creep. The other creep events (six grey arrows) were 'false alarms'.

there may be other observable phenomena, such as electromagnetic signals¹⁵ or strain events. Alternatively, the creep events may give rise to propagating creep that propagates into the seismogenic zone, leading to the initiation of the earthquakes. If aseismic creep on small patches in the seismogenic zone plays a role in the process leading up to small to moderate earthquakes (pre-slip)^{16,17}, then the shallow creep events may be useful as models for the characteristics of pre-slip. Precursory creep changes may also provide an opportunity to issue a short-term prediction for the expected magnitude 6 earthquake at Parkfield California, ~75 km southeast of this area, if the relationship reported here holds for other creeping sections and larger earthquakes. □

Received 24 November 1995; accepted 28 February 1996.

1. Wesson, R. L. *Tectonophysics* **144**, 215–229 (1987).
2. Burford, R. O. *Pur. appl. Geophys.* **126**, 499–529 (1988).
3. Wesson, R. L. *J. geophys. Res.* **93**, 8929–8951 (1988).
4. Stuart, W. D., Archuleta, R. A. & Lindh, A. G. *J. geophys. Res.* **90**, 592–604 (1985).
5. Lorenzetti, E. & Tullis, T. E. *J. geophys. Res.* **94**, 12343–12361 (1989).

6. Dieterich, J. H. in *Earthquake Source Mechanics* (eds Das, S., Boatwright, J. & Scholz, C. H.) 37–47 (American Geophysical Union, Washington, DC, 1986).
7. Ellsworth, W. L. *Bull. seism. Soc. Am.* **65**, 483–506 (1975).
8. Ellsworth, W. L. in *Urban Disaster Mitigation: The Role of Science and Technology* (eds Cheng, F. Y. & Sheu, M. A.) 1–14 (Elsevier Science, New York, 1995).
9. Schulz, S. S., Mavko, G. M., Burford, R. O. & Stuart, W. M. *J. geophys. Res.* **87**, 6977–6982 (1982).
10. Allen, C. R. & Smith, S. W. *Bull. seism. Soc. Am.* **56**, 966–967 (1966).
11. Yerkes, R. F. & Castle, R. O. *Prof. Pap. U.S. geol. Surv.* **579**, 40–52 (1967).
12. Shifflett, H. & Witbaard, R. *Eos (suppl.)* **76**, F359–F360 (1995).
13. Breckenridge, K. S. & Simpson, R. W. *Open File Rep. U.S. geol. Surv.* **95-830** (1995).
14. Minster, J.-B. H. & Williams, N. P. *Eos (suppl.)* **76**, F358 (1995).
15. Fraser-Smith, A. C., Bernardi, A., McGill, P. R., Ladd, M. E. & Helliwell, R. E. *Geophys. Res. Lett.* **17**, 1465–1468 (1990).
16. Ellsworth, W. L. & Beroza, G. C. *Science* **268**, 851–855 (1995).
17. Dodge, D. A., Beroza, G. C. & Ellsworth, W. L. *Eos (suppl.)* **76**, F388 (1995).

ACKNOWLEDGEMENTS. I thank R. Simpson for assistance with Fig. 1 and with access to the USGS creepmeter data, W. Ellsworth, P. Reasenber, Y. Ben-Zion, H. Wang, K. Breckenridge, M. Johnston, W. Stuart, R. Bilham, M. Zoback, J. Valley, R. Meyer and W. Unger for their advice and assistance, and J. Savage for comments. This work was supported by the Woollard Fund of the Geophysical and Polar Research Center, the Department of Geology and Geophysics, University of Wisconsin-Madison, and by the Continental Dynamics Program of the US NSF.

Reinterpretation of *Yunnanozoon* as the earliest known hemichordate

D. Shu, X. Zhang & L. Chen

Department of Geology, Northwest University, Xian 710069, China

THE Chengjiang fossil Lagerstätte is one of the earliest and most important palaeontological sites from the Phanerozoic era^{1,2}, about 530 million years ago³. It yields extremely abundant and remarkably preserved soft-bodied fossils and shells with soft parts of various kinds, including bradoriids^{4–6}, trilobites^{7,8}, crustaceans⁹, brachiopods, worms, sponges, algae and many unknown forms^{10–13}. One of these fossils is *Yunnanozoon*¹⁴, which we reinterpret here as the earliest known hemichordate. Possessing half of the characteristic chordate features and providing an anatomical link between invertebrates and chordates¹⁵, Hemichordata is a minor but important phylum in evolutionary biology. Hemichordates comprise two main groups: the enteropneusts, or 'acorn worms', and the pterobranchs. Apart from the presumable inclusion of graptolites in pterobranchs^{16–19}, there are very few hemichordate fossils^{2,17,20}. Although *Yunnanozoon* is superficially similar to the chordates²¹, its typical tripartite body plan is broadly consistent with that of living balanoglossid hemichordates (enteropneusts).

Unlike sessile and colonial pterobranchs, the enteropneusts are solitary and mobile, and many of them burrow in sand and mud. They comprise only 11 known worm-like marine genera, of which *Balanoglossus* is widely known. Like all other enteropneusts, *Balanoglossus* has three body parts: the proboscis, or prosome; the short collar, or mesosome; and the elongated trunk, or metasome, which is subdivided into a gill region anteriorly and a digestive region posteriorly^{15,22–24}.

Our material provides a much more complete idea of the basic architecture of *Yunnanozoon* (Figs 1–4) than does that of Chen *et al.*²¹. Several well-preserved individuals clearly show that the main body comprises three parts: a proboscis, a collar, and a trunk which can be subdivided into a branchial and a gut region. The posterior six-sevenths of the main body has a well-developed, segmented unit on its back. Because it is thin and relatively hard, and contains no organs, it is interpreted here as a peculiar sclerotized dorsal fin.

From various specimens we know that the collar is a loop-like structure with a big notch on its ventral side. The proboscis is a solid organ in front of the collar, and is egg-shaped in lateral view and tongue-shaped in ventral or dorsal view. It has been preserved protruding outside the collar (Figs 1*f,g,i* and 2*a,b*), but could 'withdraw' into or overlap onto the collar owing to variable oblique angles of burial (Figs 1*c,e,h* and 2*c,d*). A nearly complete collar was clearly seen in a specimen reported by Chen *et al.*, but its proboscis was missing, probably lost during preparation (see Fig. 1*a* in ref. 21). Occasionally, a short, dark longitudinal line can be traced in the posterior part of the proboscis, and is tentatively interpreted as the stomochord (Figs 1*g* and 2*b*). The big notch on the ventral side of the collar may indicate the position of the mouth (Fig. 1*c,f,g*).

Chen *et al.* identified seven pairs of branchial arches in the branchial region immediately behind the collar²¹. Each of the 'arches', composed of about 20 small dark rings (Fig. 3*a–c*), is of sclerotized, annulated or spiral tubes. Seven pairs of pores on the lateral surfaces in the branchial region, which correspond to the seven pairs of gill 'arches', are recognizable. That the left and right curved, annulated 'arches' are separated by a thick layer of sediments suggests that they were convex laterally and inclined anterodorsally before compaction. The ventral end of each arch is connected with the gill pore (Figs 1*d,e* and 3*b,c*), and dorsally they appear to be attached to the lateral sides of the pharynx (Fig. 4). Hence the tube-like 'arches', essentially different from the gill arches in chordates²², are best referred to as unique branchial tubes, which may be homologous to the branchial sacs in modern enteropneusts.

Posterior to the pharynx is the gut region, comprising the anterior gut and the posterior gut, which is full of faecal contents in at least four specimens. The contents in the anterior gut, the main site of digestion, are spiral in shape (Figs 1*h* and 3*d,e*), which indicates the presence of spiral valves. The spiral intestine, which made food stay longer in the gut and digestion more efficient, is independently developed in various less-advanced fish²². The spiral valves in *Yunnanozoon* are sometimes clearly visible, even without gut contents (Figs 1*b* and 3*f,g*). However, the posterior gut is a simple elongated tube, with the anus at the posterior end (Fig. 1*h*). In addition, two more structures are also visible in some specimens: the small oval structures ventral to the gut (Figs 1*d* and 3*g*), which were observed by Hou *et al.*¹⁴ and correctly interpreted as gonads by Chen *et al.*²¹; and the striking longitudinal dark lines (Fig. 1*f,i*), the nature of which is not yet known.

The pairs of gill pores and tubes in *Yunnanozoon* are a significant synapomorph, so it must be closely related to the hemichordate–chordate group^{15,22–24}. The tripartite body plan of *Yunnanozoon* is reminiscent of that of the extant hemichordate *Balanoglossus*, and the collar and the proboscis in front of the collar in *Yunnanozoon* are similar to the collar and the proboscis of *Balanoglossus*, respectively. The proposed stomochord in *Yunnanozoon* is also characteristic of the hemichordates. Moreover, not only are the paired branchial tubes connected with the corresponding branchial pores in *Yunnanozoon* evidently homologous to the gill sacs in *Balanoglossus*, but the trunk patterns composed of the branchial region and the gut region in the two animals are also nearly identical^{15,23}. The major differences between them are that *Balanoglossus* lacks the segmented unit, and has a much more well-developed proboscis than does *Yunnanozoon*. This may just reflect the results of ecological changes from the epifaunal *Yunnanozoon* to the burrower *Balanoglossus*. The segmented unit was probably used by the epifaunal *Yunnanozoon* as a dorsal fin, but is structurally different to that found in fish, and was subsequently lost in its burrowing descendants. The descendants have also gradually developed a very strong proboscis to make burrowing more efficient. Therefore, if the interpretation is correct, we may have a reconstruction of *Yunnanozoon* (Fig. 4).

Chen *et al.* pointed out that *Yunnanozoon*, which has a laterally flattened body, could be laterally undulated²¹. We believe that,

FIG. 1 More than 40 specimens of *Yunnanozoon lividum*, including juveniles and adults on 17 slabs, were collected from Ma'an shan section, 3 km north of the Maotianshan section where Chen's material was collected²¹ in Chengjiang fossil Lagerstätte from 1991 to 1995. Scale bar: 10 mm in a, b, d, f, h, i; 3 mm in c, e, g. a–d, Segmented fin attached on the dorsal side of the soft trunk (att. fin), proboscis (prob.), collar, mouth, pharynx, gonads, spiral valves (spi) in gut and gill tubes. a, More than 20 individuals together. NWU93-1401A. b, Enlargement of the far-right individual in a, a nearly complete specimen with the anterior one-seventh of the main body bending down. c, The collar with proboscis inside and the foremost part of the trunk bent backwards. NWU95-1402. d, A specimen with both ends broken. NWU93-1413. e–i, The anterior part has the blade-like dorsal fin detached from the trunk (det fin) owing to the post-mortem in f, h; the gut with spiral content (spi. con) in it in h; and the possible stomochord in f, g; detached soft body is twisted in f, h, i. e, NWU93-1403. f, Note detachment of dorsal fin producing the rupture belt (rb) in the soft trunk. NWU93-1402. g, NWU93-1402. h, Note the posterior gut immediately below the dorsal fin, leaving no room for the 'notochord'. NWU93-1406A. i, NWU93-1407A.

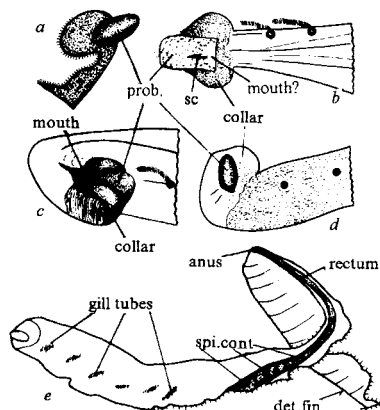
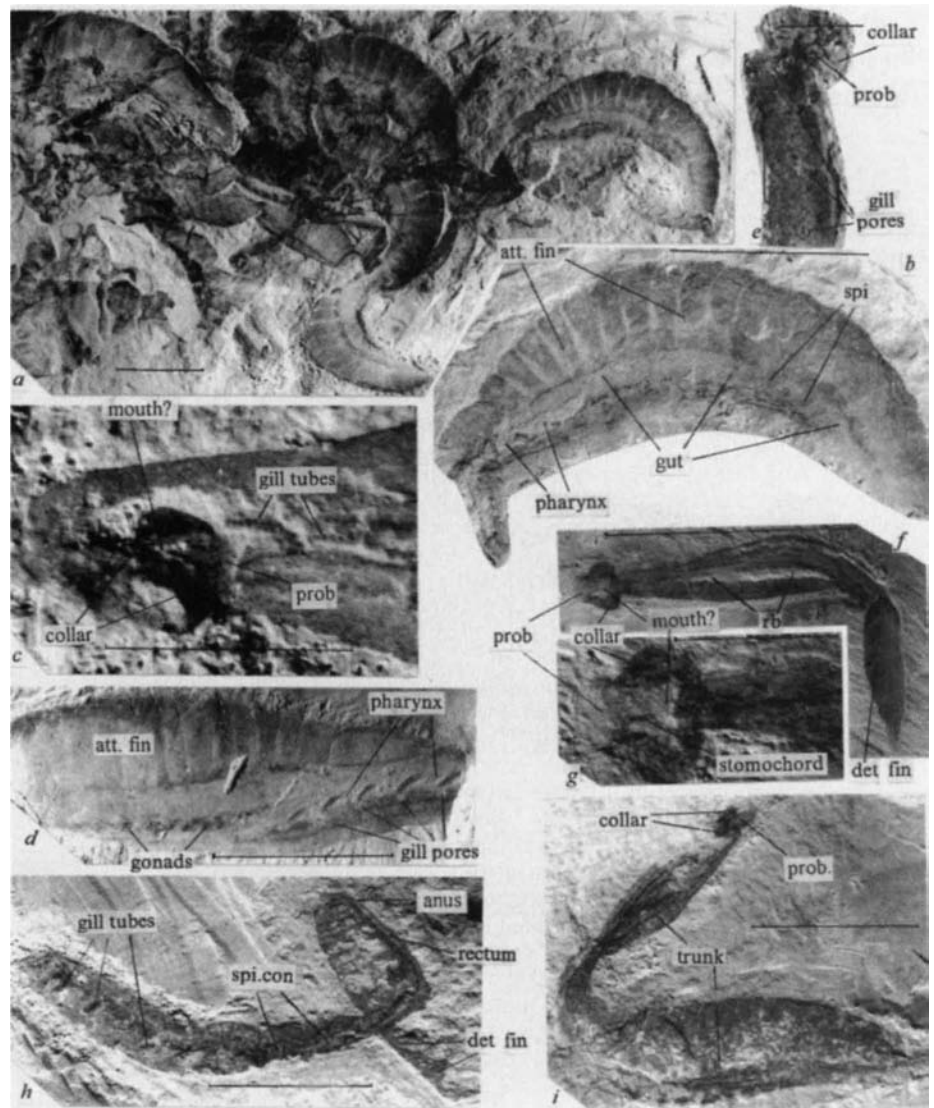


FIG. 2 Explanatory drawing of Fig. 1c, e, g, h, i, showing proboscis (prob.) with possible stomochord (sc) within it, collar and gut in detail. Note the different appearance of proboscis and collar owing to different preservation; not to scale.

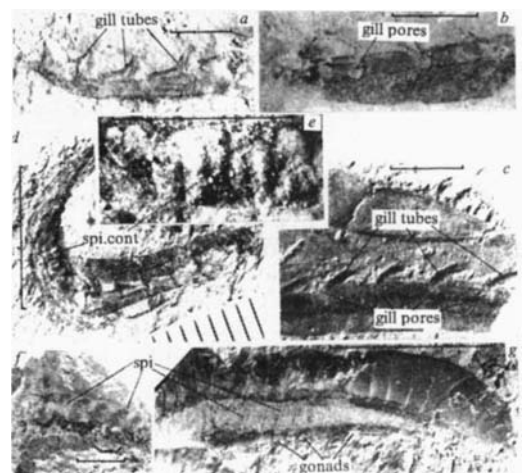


FIG. 3 a–c, Gill pores and gill tubes. a, NWU93-1401A. b, NWU93-1401B. c, NWU93-1413. d–g, Spiral contents with conspicuous relief (spi. cont.) or spiral valves (spi) in anterior gut and gonads ventral to gut. d, e, NWU93-1411. f, NWU93-1401A. g, NWU93-1416. Scale bar: 3 mm in a, b, c, e, f; 10 mm in d, g.

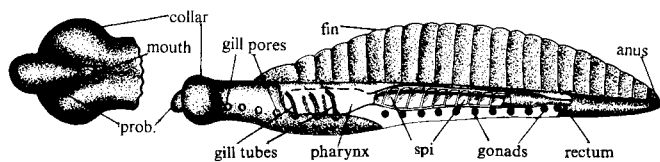


FIG. 4 Reconstruction of *Yunnanozoon lividum*, showing tripartite body plan: proboscis (prob.), collar and trunk. Also shown gill tubes, gill pores, pharynx, spiral intestine, rectum, anus, gonads, mouth and segmented fin.

with the well-developed dorsal fin, it might have been an active benthic swimmer. It would have taken water with food particles into the mouth, and the food particles would have been separated from the water in the pharynx, and transported directly to the gut. The filtered water passed through the supposed inner gill slits in the pharynx, which may not be preserved in fossils, and into the branchial tubes for oxygen exchange, before being removed through the gill pores.

Chen *et al.* assigned *Yunnanozoon* to the chordates on the basis of five features: a notochord, segmented musculature, an expanded pharynx, branchial arches, and gonads. They gave five reasons for supporting the presence of a notochord in *Yunnanozoon*, but none of them is valid. (1) They claimed that the alleged 'notochord' "is positioned dorsal to the pharynx", but it actually passes through the pharynx anteriorly as they described (ref. 21, p. 721), which should not be the case in a chordate. (2) They stated that the 'notochord' "is often straight, indicating stiffness". However, most of the 'notochord' was curved, as shown in our Figs 1a, b, h, i and 3d and their Figs 1a and 2e, and its anterior can even be angularly bent (Fig. 1b) or folded backwards (Figs 1c and 2c). (3) There seems to be no evidence for the 'notochord' to resist longitudinal compression more than other elements above and below it. (4) Without a clear shape of its anterior end being ever defined in all specimens, there is no reason for claiming that it "tapers at both ends". (5) Their 'notochord' can be interpreted as the gut, as discussed before. In addition, they described in the Fig. 2 legend that, posterior to the pharynx, "the wide, dark bands crossing the notochord . . . their nature is unknown". However, they can be easily interpreted as the spiral contents or spiral valves in the anterior gut, as shown in our Figs 1h, 2e and 3d, e, f, g. Accordingly, their alleged 'notochord' actually represents the upper pharynx and the gut following the pharynx. Furthermore, their claim that their 'notochord' is "the rod" obviously itself conflicts with their description that the "notochord is just as flattened as other soft tissues", because, according to the decay experiment²⁵, the notochord of cephalochordates is the most resistant of the soft parts to decay. Therefore, if a notochord had been present in *Yunnanozoon*, it would have been preserved as a rod-like, dark structure with conspicuous relief (rather than a pale spiral or faeces-like structure), and the trunk could have not been twisted as it was in specimen NWU93-1407 A (Fig. 1i).

A comparison of the dorsal segmented unit in *Yunnanozoon* to the myomeres in cephalochordates may also be invalid, because: (1) it is not chevron-shaped as in living cephalochordates, or sigmoidal, as in the Middle Cambrian *Pikaia*^{26,27}; (2) it is completely dorsally positioned, and never embraces visceral organs, as those in cephalochordates do; and (3) it is blade-like in shape, suggesting sclerotization, or at least half-sclerotization, and may not be recognized as musculature, especially from the detached appearance (Fig. 1f, h).

It should be stressed that although the pharynx, the branchial tubes ('arches') and the gonads could be present in chordates, the first two are crucial features in hemichordates, and the last can be present in various phyla including the Hemichordata^{15,23}. All of these facts lead to a conclusion that all of the five features, plus the important tripartite body plan, strongly support the affinities of

Yunnanozoon with the enteropneust hemichordates rather than the chordates. □

Received 7 November 1995; accepted 1 February 1996.

1. Conway Morris, S. *Nature* **361**, 219–225 (1993).
2. Chen, J., Ramsköld, L. & Zhou, G. *Science* **264**, 1034–1068 (1994).
3. Bowring, S. A. *et al.* *Science* **261**, 1293–1298 (1993).
4. Huo, S. & Shu, D. *Cambrian Bradioridia of South China* (Northwest Univ. Press, Xian, 1985).
5. Shu, D. *Cambrian and Lower Ordovician Bradioridia from Zhejiang, Hunan and Shaanxi* (Northwest Univ. Press, Xian, 1990).
6. Shu, D. *Cour. Forsch.-Inst. Senck.* **123**, 315–330 (1990).
7. Zhang, W. *Acta palaeont. sin.* **26**, 223–235 (1987).
8. Shu, D., Geyer, G., Chen, L. & Zhang, X. *Beringeria* (suppl.) **2**, 203–241.
9. Shu, D., Zhang, X. & Geyer, G. *Alcheringa* **19**, 333–342 (1996).
10. Hou, X. *Acta palaeont. sin.* **26**, 286–298 (1987).
11. Ramsköld, L. *Lethaia* **25**, 443–460 (1992).
12. Shu, D. *et al.* *NW Univ. (suppl.)* **22**, 31–38 (1992).
13. Shu, D. & Chen, L. *J. SE-Asian Sci.* **9**, 289–299 (1994).
14. Hou, X., Ramsköld, L. & Bergström, J. *Zool. Scripta* **20**, 395–411 (1991).
15. Barnes, R. S. K., Calow, P., Olive, P. J. W. & Golding, D. W. *The Invertebrates—a New Synthesis* (Blackwell Scientific, Oxford, 1988).
16. Kozłowski, R. *Palaeont. pol.* **3**, 1–235 (1949).
17. Broadman, R., Cheetham, A. & Rowell, A. *Fossil Invertebrates* (Blackwell Scientific, Oxford, 1987).
18. Bengtson, S. & Urbanek, A. *Lethaia* **19**, 293–308 (1986).
19. Durman, P. N. & Sennikov, N. V. *Palaeontology* **36**, 283–296 (1993).
20. Arduini, P., Pinna, G. & Teruzzi, G. *Atti Soc. Ital. Sci. nat. Museo civ. Storia nat. Milano* **122**, 104–108 (1981).
21. Chen, J., Dzik, J., Edgecombe, G. D., Ramsköld, L. & Zhou, G. *Nature* **377**, 720–722 (1995).
22. Romer, A. S. *The Vertebrate Body*, (W. B. Saunders, Philadelphia, 1964).
23. Barnes, R. D. *Invertebrate Zoology*, 1018–1028 (Saunders College, Philadelphia, 1980).
24. Jefferies, R. P. S. *The Ancestry of the Vertebrates*, (British Museum, London, 1986).
25. Briggs, D. E. G. & Kear, A. J. *Lethaia* **26**, 275–287 (1994).
26. Conway Morris, S. (ed.) *Atlas for the Burgess Shale* (Palaeontological Association, London, 1982).
27. Briggs, D. E. G., Erwin, D. H. & Collier, F. J. *The Fossils of the Burgess Shale*, 197–198 (Smithsonian Institution Press, Washington, DC, 1994).

ACKNOWLEDGEMENTS. We thank J. Dzik for suggestions, L. Ramsköld for discussion, and L. Li and Z. Luo for preparing the figures. This work was supported by the National Foundation of Natural Science of China.

Predicting animal cadmium concentrations in lakes

Landis Hare & André Tessier

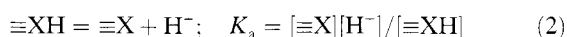
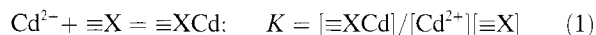
INRS-Eau, Université du Québec, Sainte-Foy, Québec G1V 4C7, Canada

HUMAN activities have greatly increased the flux of many potentially toxic metals to aquatic ecosystems¹. The development and implementation of effective remedial measures depend on our ability to predict the fate and effects of metals in these systems. Models based on sound physical-chemical and biological principles, such as the free-ion activity model^{2–5}, have shown great potential as predictive tools. This model has been effective in explaining the central role of the free-ion concentration (or activity) as a regulator of interactions (uptake, toxicity) between metals and aquatic organisms^{2,3}. It postulates that the biological effects of metals are best predicted by the activity of the free metal ion, rather than by total metal concentration. Because this model was developed in the laboratory under unnatural experimental conditions, it must be validated in field situations before being generally used in nature⁶. We report here that Cd concentrations in an indigenous aquatic insect larva, *Chaoborus punctipennis*, are best described by the free-ion activity model, provided that competition for biological uptake sites between hydrogen ions and free cadmium ions, as well as cadmium complexation by natural organic matter, are explicitly taken into account. Our results suggest that the free-ion model would provide an effective theoretical framework for the use of animals as indicators of metal contamination in nature.

Water samples and larvae of *C. punctipennis* (Say), a widespread nocturnal predator of lake plankton⁶, were collected from

23 lakes in the provinces of Ontario and Québec, Canada (Fig. 1 and Table 1). If Cd complexation with dissolved humic substances is ignored, chemical speciation computer programs predict that >94% of the aqueous Cd in the study lakes would be present as the free Cd ion (complexation by inorganic ligands such as Cl^- and SO_4^{2-} is small in these lakes). These estimates of free Cd ion concentrations ($[\text{Cd}^{2+}]^*$) were not strongly related to Cd concentrations in *C. punctipennis* larvae ($[\text{Cd}]_{\text{Chaob}}$; Fig. 2a, $r^2 = 0.13$, $P > 0.1$). The most striking aspect of the weak relationship between $[\text{Cd}^{2+}]^*$ and $[\text{Cd}]_{\text{Chaob}}$ is the position of the three most acidic lakes (pH < 5.0; filled symbols in Fig. 2a), in which Cd concentrations in larvae are low despite relatively high $[\text{Cd}^{2+}]^*$.

Considering the tenets of the free-ion model, we considered that in the more acidic lakes hydrogen ions would effectively compete with cadmium ions at biological uptake sites. It has been shown in the laboratory that, for a given free metal ion concentration, hydrogen ions reduce both the toxicity of Cd to fish⁷ and the uptake of Cu, Pb and Zn by algae⁸ and bacteria⁹. A simple representation of Cd-proton competition for a biological uptake site ($\equiv\text{X}$) is:



where charges on the sites are omitted for simplicity and K and K_a are equilibrium constants. The total concentration of uptake sites is given by:

$$[\text{X}]_T = [\equiv\text{XH}] + [\equiv\text{X}] + [\equiv\text{XCd}] \quad (3)$$

which, if combined with the expressions for the equilibrium constants in equations (1) and (2) and assuming that only a small fraction of the sites is occupied by Cd, gives

$$[\equiv\text{XCd}] = \frac{KK_a[\text{X}]_T}{[\text{H}^+] + K_a} [\text{Cd}^{2+}] \quad (4)$$

If we assume that Cd accumulated by *C. punctipennis* is proportional to $[\equiv\text{XCd}]$, that is $[\text{Cd}]_{\text{Chaob}} = k[\equiv\text{XCd}]$, combining this

TABLE 1 Concentrations of dissolved constituents and Cd in *Chaoborus*

Lake	Region	$[\text{Cd}]_{\text{Chaob}}$ ($\mu\text{g g}^{-1}$)	$[\text{Cd}]$ (nM)	$[\text{Zn}]$ (nM)	$[\text{Ca}]$ (μM)	$[\text{Mg}]$ (μM)	$[\text{C}_{\text{org}}]$ ($\text{mg l}^{-1} \text{C}$)	pH
Adeline	R.-Noranda	2.88	0.29	9	197	94	5.57	7.26
Bird	Muskoka	0.24	0.19	44	107	35	5.70	6.37
Bousquet	R.-Noranda	6.27	2.37	117	106	50	14.6	6.38
Bigwind	Muskoka	0.94	0.38	—	94	40	6.50	6.04
Caron	R.-Noranda	8.89	3.14	228	245	100	13.3	7.11
Chub	Muskoka	1.94	0.32	104	47	24	7.42	5.50
Clearwater	Sudbury	3.99	4.90	346	137	57	3.40	4.79
Crooked	Sudbury	1.43	7.12	372	77	44	2.52	4.58
Crowley	Sudbury	13.3	2.18	152	69	38	1.29	5.86
Dufresnoy	R.-Noranda	1.11	0.31	10	233	107	5.50	7.24
Dufay	R.-Noranda	1.76	0.32	16	89	71	10.6	6.55
Heva	R.-Noranda	1.88	0.67	45	52	39	12.0	6.18
Hélène	R.-Noranda	0.24	0.15	8	500	239	6.01	7.18
Joannès	R.-Noranda	2.60	1.13	42	188	80	9.26	7.27
Marlon	R.-Noranda	4.80	1.43	57	49	300	9.06	7.10
Plastic	Muskoka	1.34	0.54	123	52	21	3.74	5.94
Ril	Muskoka	0.32	0.20	89	113	35	5.68	6.02
St. Joseph	Québec	0.65	0.15	46	97	19	2.94	6.63
Sunken	Muskoka	0.72	0.25	71	52	21	6.26	6.12
Tantaré	Québec	0.35	0.40	138	47	24	4.02	5.65
Tilton	Sudbury	7.77	1.93	168	97	46	6.17	5.81
Vaudray	R.-Noranda	5.72	1.19	69	89	52	6.31	6.58
Wavy	Sudbury	2.47	2.23	223	52	30	3.07	4.62

Mean concentrations of cadmium in larvae of *Chaoborus punctipennis*, $[\text{Cd}]_{\text{Chaob}}$, and dissolved in water from the 23 study lakes. Also indicated are total dissolved concentrations of the trace metal Zn, the major ions Ca and Mg, and organic carbon, $[\text{C}_{\text{org}}]$, as well as pH. *Chaoborus punctipennis* larvae were collected during the daytime from littoral lake sediments. Live final instar larvae were held in lake water to defecate their gut contents, then pooled into 3–5 replicate samples (most lakes) of 10–20 individuals each and treated for Cd analysis¹⁷. Water was collected in dialysis samplers¹⁸, placed near the collection site for larvae, for the measurement of pH, inorganic and organic carbon, major ions (Ca, Mg, Na, K, Cl, SO_4), and trace metals (Cd, Cu, Ni, Pb, Zn); values for each lake are generally the mean of 15 measurements, that is, five samples at 1-cm intervals above the sediment–water interface from each of three dialysis samplers. Cadmium in water and animal samples was measured by flameless atomic-absorption spectrophotometry.

relation with equation (4) gives

$$[\text{Cd}]_{\text{Chaob}} = F \frac{[\text{Cd}^{2+}]}{[\text{H}^+] + K_a} \quad (5)$$

where $F (= kKK_a[\text{X}]_T)$ is a constant specific to *C. punctipennis*. As expected, prediction of Cd in larvae is substantially improved when H^+ competition with Cd^{2+} for biological uptake sites is taken into account (Fig. 2b; $r^2 = 0.66$, $P < 0.001$). The values of K_a and F , as defined in equation (5), were determined by least-squares optimization.

Although our predictive capacity is now much greater, 34% of the variability between Cd concentrations in the animals and in their environment remains unexplained. Some of this variability could be due to the influence of dissolved organic matter on Cd speciation, an interaction that has been very difficult to quantify and thus has tended to be ignored. We used the Windermere humic aqueous model (WHAM¹⁰) to address this problem. Concentrations of fulvic and humic acids required as input data to the WHAM 1.0 computer code were estimated from our measure-

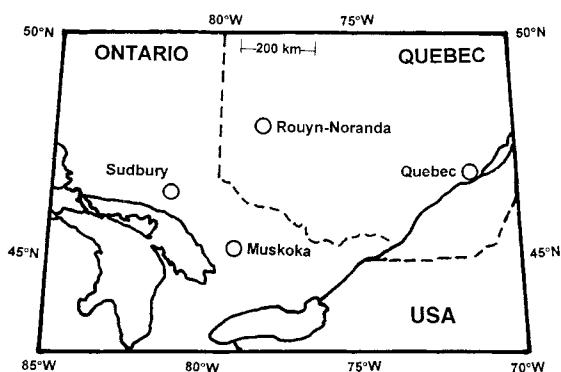


FIG. 1 Locations (○) of 23 eastern Canadian lakes in which measurements were made of cadmium in water and in larvae of the dipteran insect *Chaoborus punctipennis*. Lakes were chosen to encompass a wide range of chemical conditions (Table 1). Study lakes in the mining areas of Sudbury and Rouyn-Noranda are subject to relatively high atmospheric metal deposition from nearby smelters, whereas those in the Muskoka and Quebec City areas are far from mining activities. Insect larvae of the same developmental stage (~10 months old) were sampled from all lakes at the same time of year (springtime 1987–94) to minimize possible differences in the age and Cd-exposure history of larvae from the various lakes. The results of an intensive temporal study on $[\text{Cd}]$ in *C. punctipennis* from one of our study lakes¹⁹ suggest that $[\text{Cd}]_{\text{Chaob}}$ is fairly stable during the 6 months before our spring sampling time.

ments of dissolved organic carbon ($[C_{org}]$), by assuming that: humic substances contain 50% carbon¹¹; the ratio of humic to fulvic acids is 1:9¹²; and all dissolved organic carbon is present as humic substances (such as humic and fulvic acids). The WHAM model predicts that Cd complexation is dominated by humic substances in our study lakes, and that the free Cd ion ($[Cd^{2+}]$) represents from 17 to 88% of the total dissolved Cd present. Support for the free-ion model comes from the fact that unexplained variability in $[Cd]_{Chaob}$ among lakes is reduced by more than half when Cd complexation by dissolved organic matter is taken into account, that is, using $[Cd^{2+}]$ ($r^2 = 0.86$, $P < 0.001$; Fig. 2c) rather than $[Cd^{2+}]^*$ ($r^2 = 0.66$, $P < 0.001$; Fig. 2b).

In the WHAM model calculations, we assumed that all dissolved organic carbon was present as humic substances. However, in some lakes humic substances represent from 60 to 80% of C_{org} ^{13–15}. If we assume a value of 60% humic substances, the percentage of the explained variance of the regression shown in Fig. 2c is not altered and the values of the constants F and K_a change only slightly. Note that the relationship between $[Cd^{2+}]$

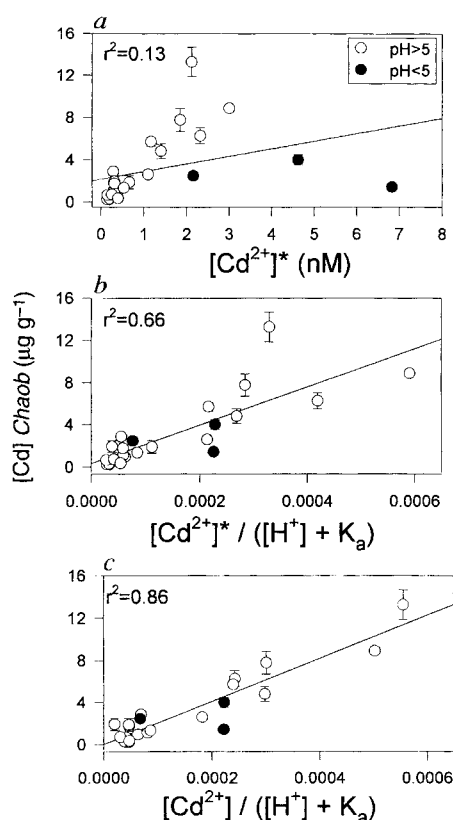


FIG. 2 Relationships between cadmium concentrations in water and in larvae of the insect *Chaoborus punctipennis* ($[Cd]_{Chaob}$) collected from 23 eastern Canadian lakes; closed symbols, pH < 5.0; open symbols, pH > 5.0. Values for $[Cd]_{Chaob}$ are means $\pm$ s.d. ($\mu\text{g g}^{-1}$ dry weight). a, Insect Cd versus computed free Cd ion concentrations if Cd complexation with humic substances is not taken into account ($[Cd^{2+}]^*$). b, Insect Cd versus $[Cd^{2+}]^*$, normalized for hypothesized competition between hydrogen and free Cd ions for biological uptake sites (see equations (1)–(5)). c, As in b except that complexation by dissolved organic substances was considered in the calculation of the free Cd ion ($[Cd^{2+}]$). The chemical speciation model WHAM 1.0, used to account for complexation with humic substances, has been validated using measurements from a range of different environments and investigators²⁰; the database (equilibrium constants, diffuse layer parameters, molecular masses of fulvic and humic acids) was not changed for our calculations of $[Cd^{2+}]$. The values ($\pm$ s.e.) of the slope ($= F$ in equation (5)) and the y intercept of the regression in c are $20,500 \pm 1,800 \mu\text{g g}^{-1}$ and $0.04 \pm 1.3 \mu\text{g g}^{-1}$, respectively, whereas $K_a = 1.9 \times 10^{-6} \text{ mol l}^{-1}$. Our estimate of K_a is close to the value ($4 \times 10^{-6} \text{ mol l}^{-1}$) reported for fish gill binding sites in a laboratory study of H^+ – Cd^{2+} competition⁷.

and $[Cd]_{Chaob}$ is not statistically significant ($r^2 = 0.03$, $P > 0.4$, regression not shown) unless Cd–proton competition is taken into account. Mention should also be made that the data point for the lake with the highest $[Cd]_{Chaob}$ (Crowley Lake; Table 1) has a strong influence on the percentage of explained variability, that is, with this data point removed the r^2 values for the regressions in Fig. 2b and c change from 0.66 to 0.79 (Fig. 2b) and 0.86 to 0.80 (Fig. 2c). There is, however, no justification for ignoring the data point corresponding to this low dissolved organic carbon lake. The influence of calcium ions as potential competitors with H^+ and Cd^{2+} for uptake sites on *C. punctipennis*¹⁶ was marginal (explained variability increased by only 2% with the inclusion of calcium terms in equation (5); modified equation not shown). A similar type of correction for possible competition between Cd^{2+} , H^+ and Zn^{2+} for Cd-uptake sites also provided only a small increase in predictive power (4%), at the cost of increased complexity (two additional adjustable parameters in equation (5)).

The predictive power of the relationship between solution variables (such as $[Cd^{2+}]$, $[H^+]$) and $[Cd]_{Chaob}$ could probably be improved further if: the concentrations of fulvic and humic acids could be measured directly rather than assuming they make up certain fixed proportions of all of the organic carbon present; and the source of Cd for *C. punctipennis* could be determined. Knowledge of the Cd source could help to determine whether the hypothesized competition between H^+ and Cd^{2+} occurs for uptake sites on the predator or for sites on organisms at lower levels in the food chain leading to *C. punctipennis*. If Cd uptake from food is important, then differences in prey types among lakes could influence $[Cd]_{Chaob}$.

Our data suggest that cadmium concentrations in an aquatic insect are modulated by both Cd complexation with organic matter as well as by competition between free cadmium and hydrogen ions for biological uptake sites. These results support the laboratory-based free-ion model as an effective tool in predicting metal concentrations in aquatic organisms in nature. A practical consequence of our study is the demonstration that larvae of *C. punctipennis* could be used as a 'sentinel' for biologically relevant cadmium concentrations in lake water, such as $[Cd^{2+}]$. Use of this species and other congeners as biomonitors would be helped by the fact that the genus has a global distribution and is often common in lakes covering a wide range of pH, organic matter and metal concentrations (as in our study). The deterministic nature of the free-ion model suggests that it should be effective in predicting the concentrations of trace metals in both freshwater and marine organisms. □

Received 7 August 1995; accepted 2 February 1996.

- Salomons, W. & Forstner, U. *Metals in the Hydrosphere* (Springer, Berlin, 1984).
- Sunda, W. G. *Biol. Oceanogr.* **6**, 411–442 (1992).
- Campbell, P. G. C. *Metal Speciation and Bioavailability in Aquatic Systems* (eds Tessier, A. & Turner, D.) 45–102 (Wiley, New York, 1995).
- Morel, F. M. M. & Hering, J. G. *Principles and Applications of Aquatic Chemistry* (Wiley, New York, 1993).
- Sunda, W. G. & Huntsman, S. A. *Limnol. Oceanogr.* **28**, 924–934 (1983).
- Yan, N. D., Nero, R. W., Keller, W. & Lasenby, D. C. *Holarctic Ecol.* **8**, 93–99 (1985).
- Cusimano, R. F., Brakke, D. F. & Chapman, G. A. *Can. J. Fish. Aquat. Sci.* **43**, 1497–1503 (1986).
- Xue, H.-B. & Sigg, L. *Water Res.* **24**, 1129–1136 (1990).
- Goncalves, M. L. S., Sigg, L., Reutlinger, M. & Stumm, W. *Sci. Total Environ.* **60**, 105–119 (1987).
- Tipping, E. *Comp. Geosci.* **20**, 973–1023 (1994).
- Buffle, J. *Complexation Reactions in Aquatic Systems: An Analytical Approach* (Ellis-Norwood, Chichester, 1988).
- Malcolm, R. L. *Humic Substances in Soil, Sediment, and Water: Geochemistry, Isolation, and Characterization* (eds Aiken, G. R. et al.) 181–209 (Wiley, New York, 1985).
- Reuter, J. H. & Perdue, G. M. *Geochim. cosmochim. Acta* **41**, 325–334 (1977).
- Malcolm, R. L. *Humic Substances in the Aquatic and Terrestrial Environment* (eds Boren, H. & Allard, B.) 369–391 (Wiley, New York, 1991).
- Boerschke, R. C., Galle, E. A., Belzile, N., Gedge, R. N. & Morris, J. R. *Can. J. Chem.* (submitted).
- Stephenson, M. & Mackie, G. L. *Aquat. Toxicol.* **15**, 53–62 (1989).
- Hare, L., Carignan, R. & Huerta-Diaz, M. A. *Limnol. Oceanogr.* **39**, 1653–1668 (1994).
- Carignan, R., Rapin, F. & Tessier, A. *Geochim. cosmochim. Acta* **49**, 2493–2497 (1985).
- Hare, L. & Campbell, P. G. C. *Freshwater Biol.* **27**, 13–27 (1992).
- Tipping, E. & Hurley, M. A. *Geochim. cosmochim. Acta* **56**, 3627–3641 (1992).

ACKNOWLEDGEMENTS. For their technical assistance, we thank M.-R. Doyon and R. Rodrigue, and for their comments on the manuscript, we thank J.-C. Audclair, P. G. C. Campbell, R. De Vitre, M. Twiss and L. A. Warren. Funding is acknowledged from Human Resources Canada, the National Science and Engineering Research Council of Canada, the Ministère de l'Environnement et de la Faune du Québec, and the US Environmental Protection Agency.

Ultraviolet vision and mate choice in zebra finches

Andrew T. D. Bennett, Innes C. Cuthill,
Julian C. Partridge & Erhard J. Maier*

School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, UK

* Institute für Zoologie, Universität Regensburg, 93040 Regensburg, Germany

SEXUAL selection is one of the most actively studied areas of evolutionary biology¹⁻³, and ever since Darwin¹ birds have been probably the most popular taxon for testing the predictions about colour variation. Humans have been used to assess 'colour', an approach which may be flawed^{4,5} as many birds see in the ultraviolet (to which humans are blind), and have at least four spectral classes of retinal cone cells (humans have only three). Here we report experiments on zebra finches which test the hypothesis that the ultraviolet waveband (300–400 nm) is used in avian mate-choice decisions. We found that the ultraviolet is used, and that it probably contributes to hue perception. This finding may have wide implications for future studies of avian sexual selection and colour, and supports one hypothesized function of avian ultraviolet vision, the role of which is largely unknown.^{4,6,7}

All experiments were conducted on zebra finches (*Taeniopygia guttata*), as numerous studies of this species have shown that preferences in choice chambers relate to real sexual preference, and 'colours' as perceived by humans affect mate choice⁸⁻¹⁷. In addition, the species has an ultraviolet-sensitive cone type (λ_{max} , the wavelength of maximum absorbance as shown by microspectrophotometry, is 350–370 nm; J. K. Bowmaker, personal communication) as well as 'blue', 'red' and 'green' cones. Its plumage reflects only moderately in the ultraviolet range (Fig. 1), suggesting that our experiments may form a conservative test of the hypothesis that ultraviolet perception is generally involved in avian mate choice. The apparatus used to test mate-choice decisions (Fig. 2) was similar in basic design to that used in numerous mate-choice experiments, although it was modified so that male and female subjects could only view each other through transparent filters (Fig. 3) mounted vertically between each

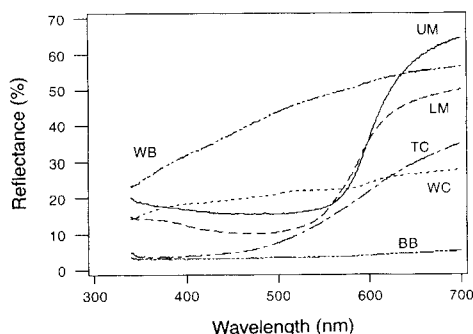


FIG. 1 Mean reflectance spectra of six body regions of five male zebra finches. Spectra are the radiance of the sample relative to that of a 99% reflection Spectralon white standard. Illumination was by xenon flash, and the reflected radiance was measured with an Oriel Intraspac IV spectroradiometer using approximately parallel radiance beams. Illumination and reflection were taken at 45° to the sample's surface. Spectra are the mean of the bird means across five randomly located measurements on each body region: UM, upper mandible; LM, lower mandible; TC, tawny cheek; WC, white cheek; BB, black bib; WB, white belly. All subjects for plumage measurements (and in experiments) were wild type, sexually experienced, and obtained, housed and fed as described previously¹³.

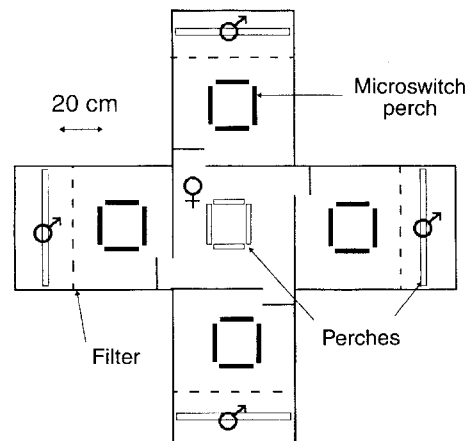


FIG. 2 Plan view of the mate-choice apparatus. The central area contains a test female; at the end of each of the four side arms is a stimulus cage containing one test male. Baffles in the central region ensure that a female can only see one male at a time. Electronically monitored perches (solid bars) placed in every side arm¹³, and linked to a data logger, allowed us to record the number of ritualized display hops^{10,17} performed by females in front of each male. Transparent filters, mounted vertically at the location of the broken lines, manipulated the wavelengths available for mate-choice decisions. Even overhead illumination was provided by 12 equispaced (at 10-cm intervals) 1.8-m, 100-W fluorescent tubes suspended 0.6 m above the test apparatus and powered by 240-V, 71-W, 35–40-kHz ballasts. These produce more ultraviolet light than normal fluorescent tubes, and have a spectral emission similar to natural skylight. Downward irradiance³ inside the stimulus cages gave $Q_{300-400\text{ nm}}/Q_{400-700\text{ nm}} = 9.2\%$ (measurements with a Spex 1681 spectrophotometer connected to an integrating sphere mounted horizontally at the centre of stimulus cages). Stimulus cages were lined with aluminium foil overlaid with frosted ultraviolet-transmitting acrylic sheet to ensure that birds were photopically adapted, and so that the horizontal radiance spectrum received by females could be estimated simply by the product of the transmission spectra of the filters (Fig. 3) and the irradiance spectra of the lights. Birds were pre-exposed to the test room and lighting conditions from 16:00 on the day before each experimental trial. During all experimental trials, perches (open rectangles), food and water were always available. Females had never previously seen the males they encountered during experimental trials.

stimulus cage and each arm. In this way, the wavelengths that were available for mate-choice decisions could be precisely manipulated. Four experiments were conducted to determine whether the ultraviolet waveband is used for avian sexual signalling, and if so, whether it is part of a hue discrimination system. For brevity, we present analyses of preference in terms of numbers of hops by females facing different stimuli; analyses of actual durations show the same patterns. Analysis was by repeated-measures analysis of variance using orthogonal contrasts to compare specific treatments¹³.

In Experiment 1 there was a different filter in each arm of the apparatus, creating the four treatments: UV+ transmitted ultraviolet and human visible wavelengths; UV− blocked ultraviolet wavelengths; ND1 and ND2 were two neutral-density filters (Fig. 3). The experiment consisted of eight trials with all birds (8 females, 32 males) being used once only. Each trial started at 10:00 and comprised three phases each of 2 h duration (Fig. 4a). Results indicated a strong preference for mate assessment to occur under UV+ conditions. As females had no preference for UV+ conditions during the control 1 and control 2 phases (Fig. 4a) when there were no stimulus males present, Experiment 1 shows that ultraviolet light is being used in mate choice and/or species recognition. To our eyes, males appeared identical when viewed through the UV+ and UV− filters; the birds were making assessments using features to which humans are blind. As there were no differences in preference for the UV−, ND1 and ND2 conditions, which only differed in the achromatic component of

light they transmitted (Fig. 3), the preference for UV+ is unlikely to be a simple preference for higher luminance³. Indeed, percentage reductions in Q_t (quantal flux³) were twice as large for UV- to ND1, and for ND1 to ND2, than for UV+ to UV- (Fig. 3), so we can be reasonably confident that we would have detected such effects. We designed and performed Experiment 2 to be more certain.

Experiment 2 was identical to Experiment 1 except that there were no control 1 and control 2 phases, and the four filter types were UV+, ND3, ND4 and ND5 (Fig. 3); UV+ was as used in Experiment 1; ND3, ND4 and ND5 were three neutral-density filters. Unlike in Experiment 1, these filters maintained their neutral density into the ultraviolet waveband. Results showed no hint of preference differences between the four filter types ($F_{3,21} = 0.56$, $P = 0.650$); neither did the two treatments with maximum ΔQ_t (UV+ and ND5) differ (paired $t_7 = 1.60$; $P = 0.150$). Experiments 1 and 2 thus support the conclusion that ultraviolet light is used for hue discrimination rather than for detecting differences in brightness.

Does the ultraviolet component of specific ornamental traits play a role in mate choice, or is 'normal' (UV+) coloration a simple prerequisite for heterosexual species recognition? Zebra finches use symmetry of both natural plumage patterns¹⁷ and artificial ornaments¹³ in mate choice. In Experiment 3 we determined whether ultraviolet asymmetry of ornaments affects preference. We manipulated the symmetry of males using leg bands, as we had done previously^{13,16}. However, here we used ultraviolet-reflecting and non-reflecting leg bands with reflectance spectra similar to the transmission spectra of the UV+ or UV- filters, respectively (Fig. 3). Each male received two UV+ and two UV- leg bands (two on each leg) in one of four arrangements, two symmetrical and two asymmetrical (see Fig. 4b and ref. 13 for details). Males were thus balanced for area of ultraviolet-reflecting and non-reflecting surfaces. On each trial the four

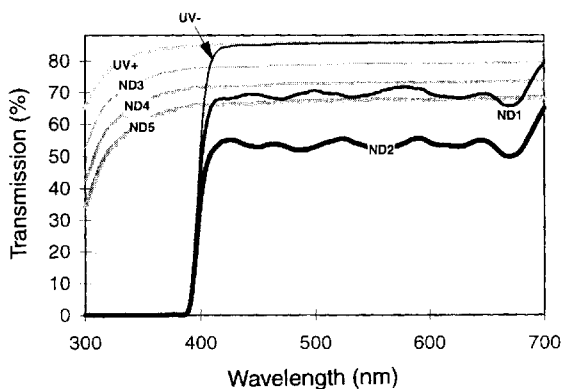


FIG. 3 Transmission spectra of the seven filter types used in the four experiments. Spectra are the mean of five randomly located measurements on each filter, taken with a Unicam Prism spectrophotometer. Filters in Experiment 1 (UV+, UV-, ND1, ND2) produced reductions in quantal flux UV+ to UV- to ND1 to ND2 in the ratio ~1:2:2 (exact reductions: 8.6, 18.8, 22.6%, respectively). These are for horizontal radiance on a female viewing stimulus cages. Filters were randomly allocated to stimulus cages, and cages to physical locations in the room, across trials. In Experiment 2, the filters used were UV+, ND3, ND4 and ND5. In contrast to Experiment 1, such filters maintained neutral density in the ultraviolet waveband. Under varying neutral density, hue remains constant but luminance varies, so if ultraviolet wavelengths are used for hue discrimination there should be no differences in preference between the four treatments. Alternatively, if ultraviolet wavelengths are used for luminance discrimination there should be a difference between treatments, and if birds maximize luminance we would expect a preference UV+ > ND3 > ND4 > ND5. In Experiment 2, the reductions in quantal flux were similar to those of Experiment 1, in the ratio 1:1:1 (exact reductions: 7.8, 7.8, 7.8%). In all other respects, Experiment 2 was identical to Experiment 1. In Experiments 3 and 4, only the UV+ filter type was used.

treatments were allocated at random to four males. Different birds were again used in each of the eight trials. All filters were UV+; in all other respects, Experiment 3 was identical to Experiment 2. Results indicate a significant preference for symmetrical males (Fig. 4b), as in previous studies^{13,16,17}. As birds could only have detected the asymmetry using ultraviolet light (and again, we could not see any differences), the experiment indicates that birds use ultraviolet light in their mate-choice decisions. Although species recognition is often inextricably linked with mate choice¹⁸, the subtle manipulations of a single novel male trait (here band asymmetry) demonstrates that species recognition cannot alone explain the use of ultraviolet light. As filters in the four arms were the same (all UV+), choosing females appeared identical to each of the four stimulus males. We can therefore also be confident that our results are due to female choice.

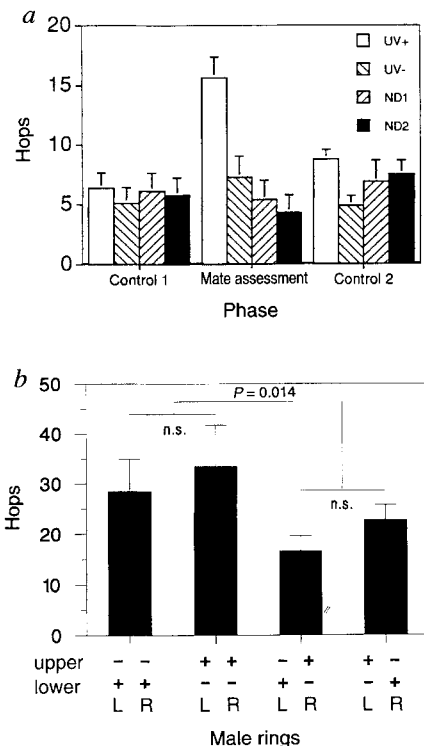


FIG. 4 a, Mean hops (+ s.e.) by females facing stimulus cages fronted by filters of different spectral transmission (Fig. 3) during the three phases of experiment 1. Phases were: 'Control 1', during which only the test female was present in the central area; 'Mate assessment', during which one male was present in each of the four stimulus cages and the female was present in the central area; and 'Control 2', during which the stimulus cages were again empty. Controls 1 and 2 tested for a preference for a particular 'coloured view' irrespective of mate choice. The distribution of activity between stimulus cages differed markedly between phases (filter-phase interaction, $F_{6,42} = 6.63$, $P < 0.001$). During control phases 1 and 2 there were no differences between treatments ($F_{3,21} = 0.64$, $P = 0.601$; and $F_{3,21} = 2.78$, $P = 0.066$, respectively), but during the mate-assessment phase treatment was highly significant ($F_{3,21} = 13.17$, $P < 0.001$). This was because there were more hops facing the UV+ males, with no differences between the other three treatments (partitioning the factor filter into two orthogonal contrasts: between UV+ and other three treatments, $F_{1,21} = 37.08$, $P < 10^{-3}$; between the other three filters, $F_{2,21} = 1.21$, $P = 0.317$). b, Mean hops (+ s.e.) by females facing stimulus cages containing males with symmetrical or asymmetrical leg bands during Experiment 3. All males received one UV+ and one UV- band on each leg; the four treatments differed only in the bands' relative positions (abscissa legend depicts the band positions on left (L) and right (R) legs). Two of the four leg-band arrangements were asymmetrical¹³, two were symmetrical¹³. Symmetrical males were preferred over asymmetrical males ($F_{1,7} = 10.73$, $P = 0.014$), with no difference within symmetry categories ($F_{2,14} = 0.32$, $P = 0.728$).

Experiment 4 was designed to confirm that our apparatus measures heterosexual mate preferences rather than, for example, the tendency to flock with individuals regardless of sex. It was identical to Experiment 2 except that: (1) two of the four stimulus cages contained females instead of males; and (2) only the UV+ filter type was used (Fig. 3). Results indicated a strong preference for associating with individuals of the opposite sex (effects of factor sex, $F_{1,7} = 139.59$, $P < 10^{-5}$); females performed 6 times as many hops, and spent 10 times longer near males than females. The experiment thus shows that our mate-choice apparatus measures heterosexual preferences, which is consistent with independent research⁸⁻¹⁶.

Taken together, our four experiments show that ultraviolet wavelengths, to which humans are blind, and hues with an ultraviolet component are used in avian mate-choice decisions. For over 20 years it has been known that birds see in the ultraviolet waveband^{4-6,19-22}. Although previous studies have postulated a role for ultraviolet light in avian^{4-6,22-27} (and reptilian²⁸) mate choice, we believe this to be the first study experimentally demonstrating such effects in any vertebrate. Birds have also been shown to use ultraviolet light for social signalling²⁷ and when hunting for rodent prey²⁹. However, it is unknown whether ultraviolet light is used in these tasks for hue discrimination or simply for perception of achromatic brightness differences. Wavelength-dependent properties of light suggest there are plausible reasons for and against ultraviolet being a 'special' waveband for avian sexual signalling⁴. Further experiments will be required to test this hypothesis; we make no claims in support of the hypothesis here and have focused on the ultraviolet waveband simply because it has been ignored in virtually all previous studies of avian sexual selection⁵. It is accepted that birdsong cannot be understood using

human hearing, or avian paternity patterns using human observation of copulation attempts. To understand the use of colour by birds, similar shifts in methodology may be required. At the very least, our experiments suggest that in future it may be unwise to ignore the ultraviolet waveband. □

Received 23 October 1995; accepted 1 February 1996.

1. Darwin, C. *The Descent of Man and Selection in Relation to Sex* (Murray, London, 1871).
2. Endler, J. A. *Phil. Trans. R. Soc. Lond. B* **340**, 215-225 (1993).
3. Endler, J. A. *Biol. J. Linn. Soc. Lond.* **41**, 315-352 (1990).
4. Bennett, A. T. D. & Cuthill, I. C. *Vision Res.* **34**, 1471-1478 (1994).
5. Bennett, A. T. D., Cuthill, I. C. & Norris, K. J. *Am. Nat.* **144**, 848-860 (1994).
6. Jacobs, G. H. *Am. Zool.* **32**, 544-554 (1992).
7. Goldsmith, T. H. *Vision Res.* **34**, 1479-1487 (1994).
8. Burley, N. *Am. Nat.* **127**, 415-455 (1985).
9. Burley, N. *Evolution* **40**, 1191-1206 (1986).
10. Burley, N., Krantzberg, G. & Radman, P. *Anim. Behav.* **30**, 444-455 (1982).
11. Burley, N. & Coopersmith, C. B. *Ethology* **76**, 133-151 (1987).
12. Collins, S. A., Hubbard, C. & Houtman, A. M. *Behav. Ecol. Sociobiol.* **35**, 21-25 (1994).
13. Swaddle, J. P. & Cuthill, I. C. *Nature* **367**, 165-166 (1994).
14. Vos, D. R. *Anim. Behav.* **50**, 645-653 (1995).
15. Clayton, N. S. *Anim. Behav.* **39**, 474-486 (1990).
16. Swaddle, J. P. *Anim. Behav.* **51**, 203-210 (1996).
17. Swaddle, J. P. & Cuthill, I. C. *Proc. R. Soc. Lond. B* **258**, 267-271 (1994).
18. Owens, I. P. F. & Short, R. V. *Trends Ecol. Evol.* **10**, 44-47 (1995).
19. Huth, H. H. & Burkhardt, D. *Naturwissenschaften* **9**, 738-740 (1972).
20. Maier, E. J. *J. comp. Physiol. A* **170**, 709-714 (1992).
21. Maier, E. J. & Bowmaker, J. K. *J. comp. Physiol. A* **172**, 295-301 (1993).
22. Finger, E. & Burkhardt, D. *Vision Res.* **34**, 1509-1514 (1994).
23. Burkhardt, D. *Naturwissenschaften* **69**, 153-157 (1982).
24. Burkhardt, D. *J. comp. Physiol. A* **164**, 787-796 (1989).
25. Burkhardt, D. & Finger, E. *Naturwissenschaften* **78**, 279-280 (1991).
26. Radwan, J. *Acta ornithol.* **27**, 125-130 (1993).
27. Maier, E. J. *Naturwissenschaften* **80**, 476-478 (1994).
28. Fleishman, L. J., Loews, E. R. & Leal, M. *Nature* **365**, 397 (1993).
29. Viitala, J., Korpimäki, E., Palokangas, P. & Koivula, M. *Nature* **373**, 425-427 (1995).

ACKNOWLEDGEMENTS. We thank J. K. Bowmaker, R. H. Douglas, J. A. Endler, D. Osorio and A. Pomiankowski for suggestions, and D. Burkhardt, E. Finger, A. R. Goldsmith, M. F. Land, K. Lunar, N. J. Marshall, J. P. Swaddle, S. Iles, R. Massie, M. Donnelly, P. Chappell and T. Colborn for discussions and assistance. This research was funded by the BBSRC.

Abnormal blood vessel development and lethality in embryos lacking a single VEGF allele

Peter Carmeliet*, Valérie Ferreira*, Georg Breier†, Saskia Pollefeyt*, Lena Kieckens*, Marina Gertsenstein‡, Michaela Fahrig‡, Ann Vandenhoeck*, Kendraprasad Harpal‡, Carmen Eberhardt†, Cathérine Declercq*, Judy Pawling‡, Lieve Moons*, Désiré Collen*, Werner Risau† & Andras Nagy‡§

* Center for Transgene Technology and Gene Therapy, Flanders Interuniversity Institute for Biotechnology, KU Leuven, B-3000 Leuven, Belgium

† Max-Planck-Institute für physiologische und Klinische Forschung, W. G. Kerckhoff-Institut, Abteilung Molekulare Zellbiologie, Bad Nauheim, Germany

‡ Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada

§ Department of Medical Genetics, University of Toronto, Ontario, Canada

THE endothelial cell-specific vascular endothelial growth factor (VEGF)¹⁻⁵ and its cellular receptors Flt-1 (refs 6,7) and Flk-1 (refs 8,9) have been implicated in the formation of the embryonic vasculature. This is suggested by their colocalized expression during embryogenesis^{10,11} and the impaired vessel formation in Flk-1 (ref. 12) and Flt-1 (ref. 13) deficient embryos. However, because Flt-1 also binds placental growth factor^{14,15}, a VEGF

homologue, the precise role of VEGF was unknown. Here we report that formation of blood vessels was abnormal, but not abolished, in heterozygous VEGF-deficient (*VEGF*^{+/-}) embryos, generated by aggregation of embryonic stem (ES) cells with tetraploid embryos (T-ES)^{16,17}, and even more impaired in homozygous VEGF-deficient (*VEGF*^{-/-}) T-ES embryos, resulting in death at mid-gestation. Similar phenotypes were observed in *F1-VEGF*^{+/-} embryos, generated by germline transmission. We believe that this heterozygous lethal phenotype, which differs from the homozygous lethality in VEGF-receptor-deficient embryos, is unprecedented for a targeted autosomal gene inactivation, and is indicative of a tight dose-dependent regulation of embryonic vessel development by VEGF.

Targeted inactivation of one (*VEGF*^{-/-}) or both (*VEGF*^{-/-}) alleles in ES cells was accomplished by replacement of the third common VEGF exon with the gene encoding neomycin phosphotransferase (neo), which caused a frameshift in the VEGF coding sequence¹⁸ (unpublished observations), and deleted six of the eight essential cysteine residues^{19,20} (Fig. 1a, b). Absence of exon 3 was confirmed by reverse transcription-polymerase chain reaction (RT-PCR) (Fig. 1c). Northern blot (Fig. 1d-g) and RT-PCR analysis (see Supplementary Information) revealed the presence of two aberrant VEGF transcripts at very low abundance (< 3% of wild type) in *VEGF*^{-/-} embryos and ES cells, differentiated to cystic embryoid bodies (CEBs), which resulted from splicing into (mRNA-1) or around *neo* (mRNA-2) and contained a stop codon after 49 residues in *neo* (mRNA-1) or after 24 out-of-frame residues in VEGF exon 4 (mRNA-2) (Fig. 1a and see Supplementary Information). These mutant transcripts were similarly detected in *VEGF*^{+/-} and in *hVEGF*^{+/-} (containing a randomly integrated targeting vector) CEBs and embryos, and resulted from residual transcription of the randomly integrated targeting vector. Transient transfection of mRNA-2 (Fig. 1a) in COS cells under the control of the cytomegalovirus promoter did not reveal

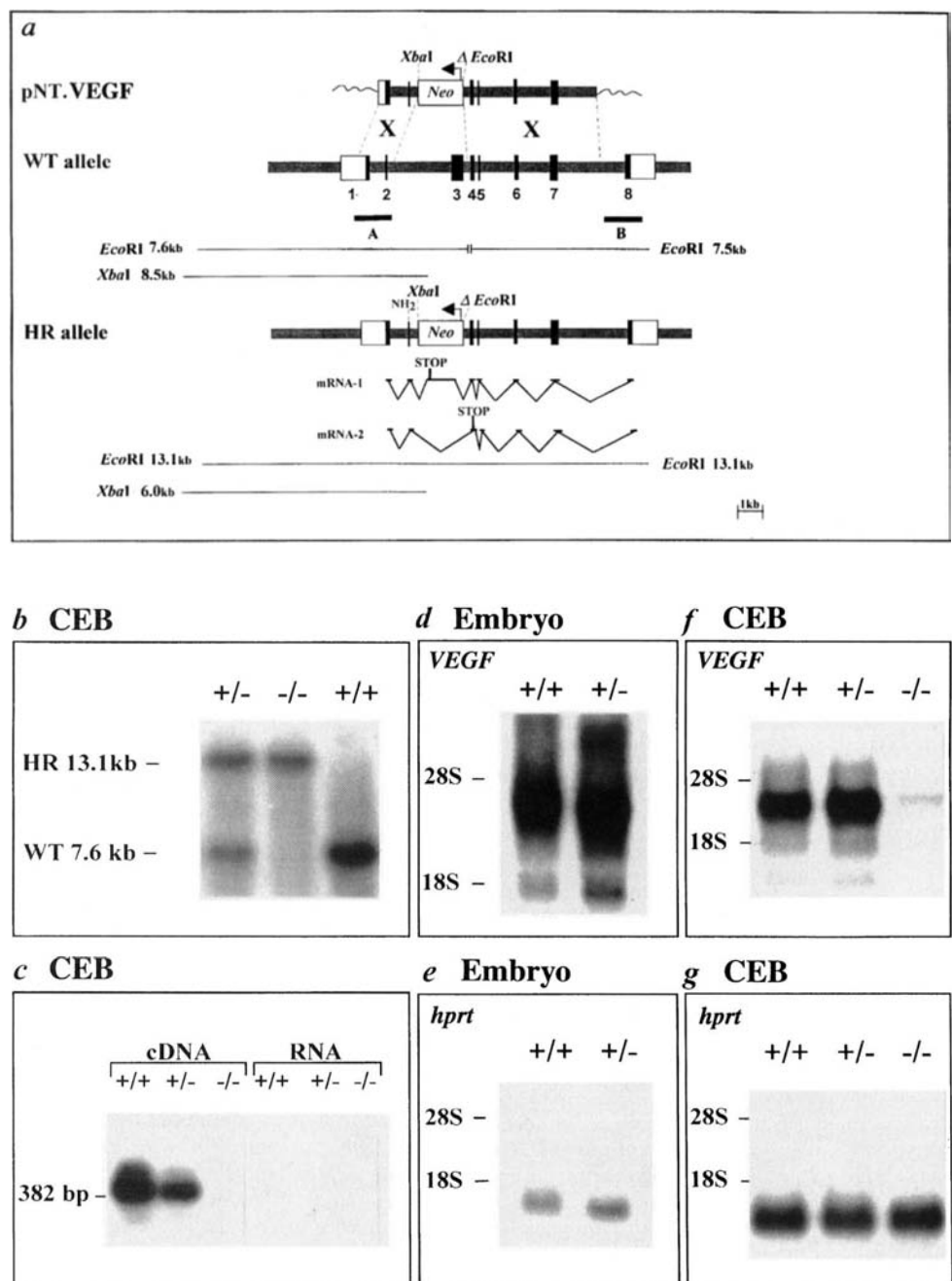
detectable expression of a mutant VEGF peptide, nor dominant-negative inhibition by such peptide on the expression level of wild-type *VEGF* in co-transfection experiments, as evaluated by western blot analysis, metabolic labelling, immunoprecipitation (for data, see Supplementary Information) and by induction of procoagulant activity (not shown), indicating that expression of mRNA-2 does not interfere with wild-type *VEGF* expression.

To distinguish between tetraploid-derived and targeted ES cell-derived embryonic cells, ROSA-26 embryos²¹, which ubiquitously express *lacZ*, were used as donors for the tetraploid embryos. The marker confirmed that tetraploid cells were confined to the endoderm of the yolk sac and trophoblast cells in all cases studied (see *LacZ*-positive cells in the yolk sac in Fig. 2d, e). *nVEGF*^{+/+}, *VEGF*^{+/-} and *VEGF*^{-/-} T-ES-cell-derived embryos appeared macroscopically normal at 8.5 days post-coitum (d.p.c.), whereas most of the *VEGF*^{+/-} and *VEGF*^{-/-} T-ES-cell-derived embryos were retarded at 9.5 d.p.c. and appeared dead at 10.5 d.p.c. At 8.5

and 9.5 d.p.c., the dorsal aorta was completely normal in the anterior (Fig. 2a) and posterior (not shown) part of *nVEGF*^{+/+} T-ES-cell-derived embryos, but was only poorly developed in the *VEGF*^{+/-} T-ES-cell-derived embryos, where endothelial cells lined a much smaller lumen than normal, especially in the anterior part of the embryo (Fig. 2b, c). At 8.5 d.p.c., the dorsal aorta was missing over its entire length in *VEGF*^{-/-} T-ES-cell-derived embryos (Fig. 2e). At 9.5 d.p.c., abnormally enlarged vascular structures, lined by endothelial cells and filled with nucleated blood cells, were observed throughout the anterior part of the necrotic embryo instead of the normal dorsal aorta and other blood vessels in *VEGF*^{-/-} T-ES-cell-derived embryos (Fig. 2f). In the caudal part of the embryo, the dorsal aorta was significantly smaller (not shown). A higher degree of tissue necrosis was observed in the *VEGF*^{-/-} than in *VEGF*^{+/-} T-ES cell-derived embryos. RT-PCR analysis for the endothelial cell-specific markers Flt-1, Flk-1 and Tie-2 (also known as Tek) (see

FIG. 1 a, Schematic representation of the targeting vector pNT.VEGF, the wild-type (WT) *VEGF* allele, and the homologously recombined (HR) *VEGF* allele. A 1.8-kb *NotI*-*Bam*HI fragment and a 5.7-kb *Eco*RI-*Nco*I fragment of the murine *VEGF* gene were cloned into pNT²⁵, which destroyed the *Eco*RI site (Δ *Eco*RI). mRNA-1 resulted from cryptic splicing from the *VEGF* exon 2 splice donor site to a splice acceptor site on the non-coding strand of *neo* (nucleotide 1,373 in pNT²⁵), followed by splicing from a donor sequence in *neo* (nucleotide 874 in pNT²⁵) to the splice acceptor site of exon 4. mRNA-2 resulted from splicing from *VEGF* exon 2, around *neo*, to exon 4. Shaded boxes represent introns; open boxes, 5' and 3' untranslated regions; filled boxes, coding regions; and closed boxes beneath the gene, hybridization probes. b, Southern blot analysis of genomic DNA, digested with *Eco*RI and hybridized to probe A, revealing the 7.6-kb and 13.1-kb *Eco*RI fragment of the WT and HR *VEGF* alleles, respectively. c, Southern blot analysis of RT-PCR products, using exon 3 and exon 8 primers for amplification and an exon 4 primer for hybridization, revealing allele dosage-dependent reduction of the expected 382-bp transcript in *VEGF*^{-/-} CEBs and absence in *VEGF*^{-/-} CEBs. Non-transcribed RNA was used as negative control. d-g, Northern blot analysis using a *VEGF* or *hprt* probe, revealing a predominant 3.8-kb mRNA in *VEGF*^{-/-} and *VEGF*^{+/-} embryos and CEBs, and a 4.1-kb mRNA in *VEGF*^{+/+} CEBs which represents mRNA-1, as described in a.

METHODS. Four different ES cell lines (D₃, JH₁, RW₄ and R₁) were used for electroporation. Culture and analysis of singly and doubly targeted ES cells were as previously described^{26,27}. RT-PCR was done as previously described²⁸, using oligomers VEGFD in exon 3 (5'-GCCCTGGAGTGCGTGCC-CACGTGAGAGAGCA-3') and VEGF8 in exon 8 (5'-GAATTCACCGCCTCGGCTT-GTCCACT-3') for amplification and oligomer VEGF11 in exon 4 (5'-CTGG-CTTTGGTGAGGTTTATCCGATG-3') for hybridization.



Supplementary Information), and *in situ* hybridization for Flt-1 and Flk-1, and immunostaining for Flk-1 and PECAM/CD31 (not shown) at 8.5 and 9.5 d.p.c. revealed fewer and frequently isolated, disorganized and scattered Flk-1- and Flt-1-expressing cells throughout the *VEGF*^{-/-} embryo at 8.5 d.p.c., and relatively low amounts of Flt-1 and Tie-2 message (which appear later during development) in *VEGF*^{-/-} embryos at 9.5 d.p.c., suggesting delayed but not aborted endothelial cell development. Thus, at both 8.5 and 9.5 d.p.c., the *VEGF*^{-/-} genotype resulted in a more severe phenotype than did the *VEGF*^{+/-} genotype.

No viable F₁ *VEGF*^{+/-} offspring were obtained at birth, indicating embryonic lethality. F₁ *VEGF*^{+/-} and F₁ *VEGF*^{+/-} : *tie-1-lacZ*^{+/-} embryos, in which the endothelial cells express *lacZ*

(owing to its targeting to the endothelial cell-specific *tie-1* locus²²), displayed macroscopic and microscopic abnormalities in embryonic development and vascular growth identical to those of the *VEGF*^{-/-} T-ES cell-derived embryos (Fig. 2g–l). Quantitative morphometric analysis revealed that fewer endothelial cells lined the smaller lumen of the dorsal aorta in the anterior part of F₁ *VEGF*^{+/-} than F₁ *VEGF*^{+/-} embryos (Table 1a). Immature development and defective interconnections of the embryonic vasculature in F₁ *VEGF*^{+/-} embryos was confirmed by *LacZ* staining of whole-mount F₁ *VEGF*^{+/-} : *tie-1-lacZ*^{+/-} embryos, which revealed a decreased density of sprouting intersomitic and head mesenchyme vessels (Fig. 3a, b), and by injection of ink in the heart, which made visible the dorsal aorta and head

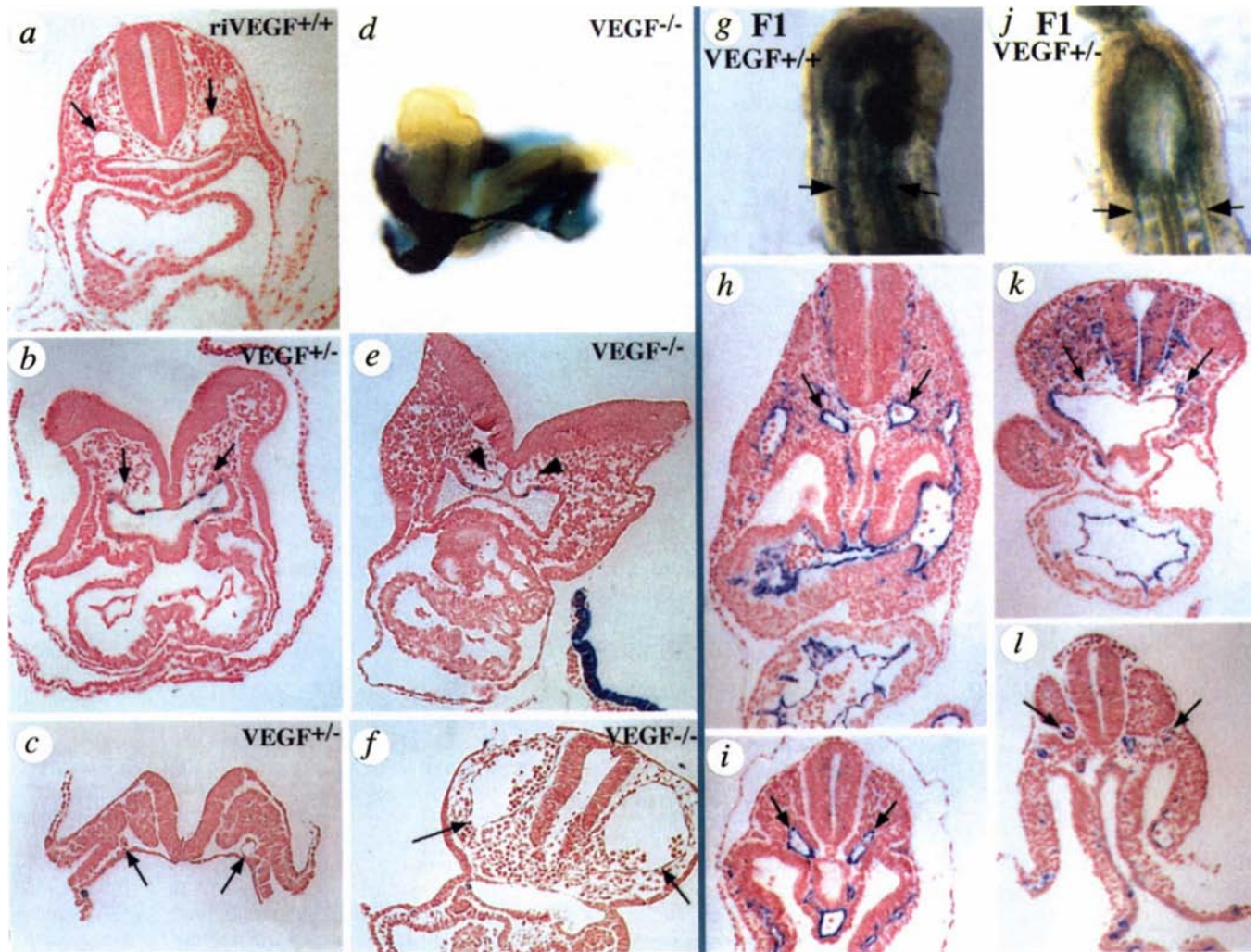


FIG. 2 Vascular development in T-ES cell-derived (a–f) and F₁ *VEGF* embryos (g–l). a, Normal vascular development in the anterior part of an 8.5-d.p.c. *riVEGF*^{+/-} T-ES cell-derived embryo. b, c, Reduced diameter of the dorsal aorta (arrows) in the anterior (b) and the posterior (c) part of an 8.5-d.p.c. *VEGF*^{-/-} embryo. d, LacZ staining of a whole-mount 8.5-d.p.c. *VEGF*^{-/-} embryo, revealing the yolk sac (blue, tetraploid embryo-derived) and the unstained embryo proper (ES-cell-derived). e, Absence of the dorsal aorta in the anterior part of an 8.5-d.p.c. *VEGF*^{-/-} embryo. Arrowheads indicate where the dorsal aorta normally develops. f, Abnormally enlarged vascular structures (arrows in f) and significant tissue necrosis in the anterior part of a 9.5-d.p.c. *VEGF*^{-/-} embryo. g, j, Whole-mount LacZ staining of the posterior part of a 9.5-d.p.c. F₁ *VEGF*^{+/-} : *tie-1-lacZ*^{+/-} (g), revealing a normal dorsal aorta and of a 9.5-d.p.c. F₁ *VEGF*^{+/-} : *tie-1-lacZ*^{+/-} (j) embryo, revealing the smaller diameter of the dorsal aorta (arrows point to the dorsal aorta). h, i, Normal dorsal aorta development in both the anterior (h) and posterior (i) part of a 9.5-d.p.c. wild-type embryo.

k, l, The dorsal aorta in the 9.5-d.p.c. F₁ *VEGF*^{+/-} : *tie-1-lacZ*^{+/-} embryo was poorly developed in the anterior part, where only a few blue cells formed a string-like aorta (k), and somewhat better in the posterior part, where the dorsal aorta contained a small lumen (l).

METHODS. Completely ES cell-derived embryos were made by aggregation with tetraploid embryos^{16,17}, using *riVEGF*^{+/-} and *VEGF*^{-/-} clones and three daughter *VEGF*^{-/-} clones that survived the high G418 selection without disruption of the second *VEGF* allele and four daughter *VEGF*^{-/-} clones. No phenotypic differences were observed between embryos obtained from these *VEGF*^{-/-} clones. *VEGF*:*tie-1-lacZ* embryos were genotyped by PCR on yolk sac-derived DNA, using primers *VEGFD* and *VEGF8* (see Fig. 1 legend), and *neo*-specific primers *neoA* (5'-ATTGAACAAGATGGATTGCAC-3', sense direction) and *neoB* (5'-TTCGTCCAGATCATCTGATCGAC-3', antisense direction). Embryos and yolk sac were dissected, fixed in glutaraldehyde, and stained for *lacZ* overnight^{12,17}, then postfixed in 4% formalin, paraffin embedded and stained with haematoxylin–eosin.

TABLE 1 Morphometric analysis of blood-vessel development in 9.5-d.p.c. F_1 $VEGF^{+/-}$ and F_1 $VEGF^{-/-}$ embryos

	F_1 $VEGF^{+/-}$	F_1 $VEGF^{-/-}$
(a) Embryo		
Endothelial cells in the dorsal aorta		
Anterior	14 $\pm$ 0.7 (5)	5.0 $\pm$ 0.6 (5) ***
Posterior	12 $\pm$ 1.5 (4)	8.0 $\pm$ 1.0 (7)
(b) Yolk sac		
Sections with large collecting vessels	73/104 (74%)	0/54 (0%)‡***
Size of small vessels (μ m)†	90 $\pm$ 10	150 $\pm$ 20§***
Ratio of the cross-sectional vessel/yolk sac length (vessel density)†	0.82 $\pm$ 0.01	0.62 $\pm$ 0.07 *
Number of endoderm cells surrounding a small vessel†	8.4 $\pm$ 0.5	15.9 $\pm$ 2.8 *
Number of endothelial cells lining the endodermal side of a small vessel†	1.2 $\pm$ 0.2	2.7 $\pm$ 0.4 *
Ratio of endothelial/endoderm cells per small vessel (endothelial cell density)†	0.15 $\pm$ 0.02	0.17 $\pm$ 0.02
BrdU-immunoreactive endothelial cells lining the endodermal side of a small vessel	159/428 (37%)	144/398 (36%)‡

(a) The data represent the mean $\pm$ s.e.m. of the number of endothelial cells lining the dorsal aorta (3–5 cross-sections per embryo), with the number of embryos indicated in parentheses. (b) The percentage of sections containing at least one large collecting yolk-sac vessel (arrow in Fig. 3h) was determined in four yolk sacs. The size (mean $\pm$ s.e.m.) of the small yolk-sac vessels is defined as the cross-sectional length, adjacent and parallel to the endoderm layer (c in Fig. 3g). Vessel density was defined as the ratio of the sum of the total cross-sectional vessel length, adjacent to the endodermal layer (c in Fig. 3g) versus total yolk-sac length (t in Fig. 3g) analysed (mean $\pm$ s.e.m.). For the number of endoderm cells (mean $\pm$ s.e.m.) surrounding a small vessel, see arrow in Fig. 3g. For the number of endothelial cells (mean $\pm$ s.e.m.) lining the endodermal side of the vessel, see arrowhead in Fig. 3g. Endothelial cell density was defined as the ratio of endothelial versus endoderm cells per vessel (mean $\pm$ s.e.m.). Cell counts represent the number of nuclear profiles. 5'-Bromo-2'-deoxyuridine (BrdU) labelling of cells was done by incubating yolk sacs in medium containing 30 μ M BrdU for 60 min at 37 °C, followed by fixation in 4% formalin. BrdU immunostaining was performed using a biotinylated rat anti-BrdU antibody (clone BU1/75; Sera Lab, Sussex, UK) in an indirect immunostaining procedure, as previously described²⁵.

† Measurements were taken using 80 yolk-sac vessels in four different yolk sacs.

‡ Statistical analysis used χ^2 tests.

§ Statistical analysis used Mann-Whitney test.

|| Statistical analysis used analysis of variance.

* $P < 0.05$; *** $P < 0.001$ versus F_1 $VEGF^{-/+}$.

vessels in $VEGF^{+/-}$ T-ES cell-derived embryos but revealed no connection between the heart and the vessel system in $VEGF^{+/-}$ T-ES cell-derived embryos (Fig. 3c, d).

The formation of the extraembryonic vasculature in 9.5-d.p.c. F_1 $VEGF^{+/-}$ embryos was also defective. In addition to a dense plexus of small vessels, larger collecting vessels that connect with the embryo were observed in the yolk sac of 9.5-d.p.c. F_1 $VEGF^{+/-}$ embryos (Fig. 3e, h). In contrast, the yolk sac of F_1 $VEGF^{+/-}$ embryos displayed an irregular plexus of small vessels but no larger collecting vessels (Fig. 3f, i). These results were confirmed by whole-mount staining with PECAM-specific antibodies of the yolk-sac vasculature (not shown) and by quantitative morphometric analysis, indicating the absence of such large collecting vessels in F_1 $VEGF^{+/-}$ yolk sacs (Table 1b). Morphometric analysis further showed that the diameter of the small vessels was enlarged but that fewer small vessels were present in F_1 $VEGF^{+/-}$ yolk sacs compared to those in F_1 $VEGF^{+/-}$ yolk sacs. The endothelial cell density and proliferation rate were, however, similar in both genotypes (Table 1b). This may indicate that VEGF deficiency in the yolk sac affects normal fusion of angioblasts and the spatial organization of endothelial cells.

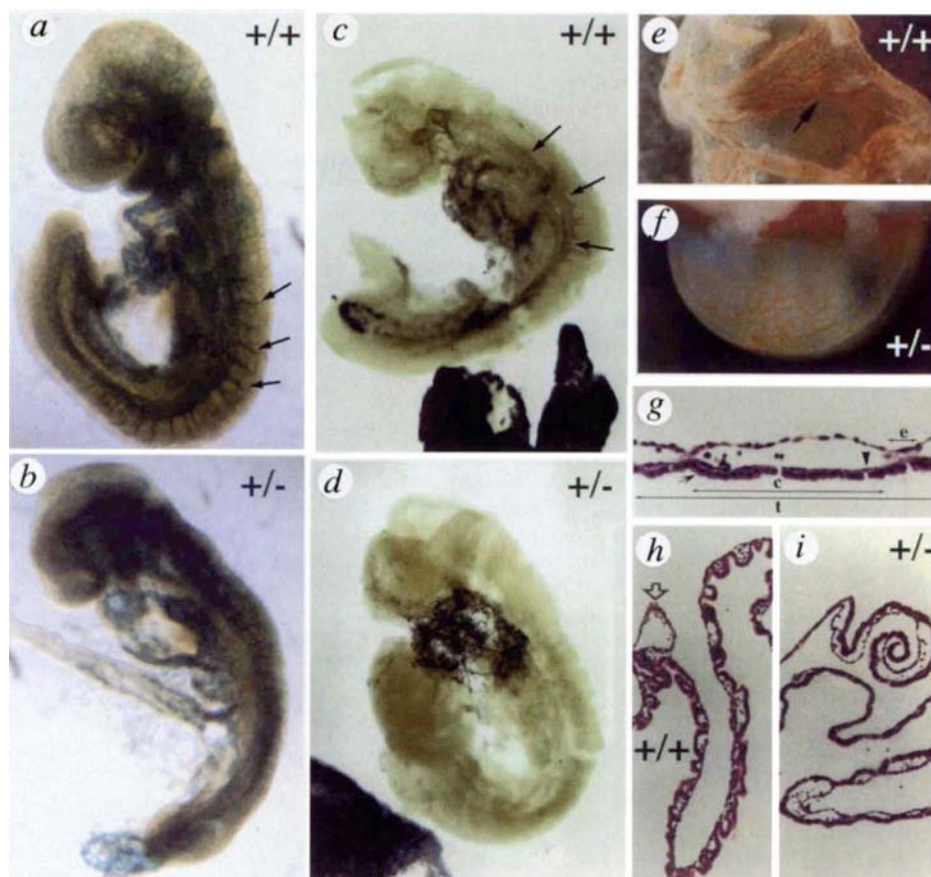
This study demonstrates that VEGF deficiency impaired most steps of early vascular development²³, including *in situ* differentiation of blood islands (vasculogenesis), sprouting from pre-existing vessels (angiogenesis), lumen formation, the formation of large vessels, the establishment of interconnections, and the spatial organization of intra- and extra-embryonic vessels. Several factors may have contributed to this phenotype, including abnormal accumulation or delayed differentiation of endothelial cells (by VEGF acting as an endothelial cell-specific growth factor²⁻⁵), irregular fusion of angioblasts, and network formation of endothelial cells (possibly by VEGF acting in mediating cell–cell or cell–matrix contacts or in providing positional information), or by accelerated regression of pre-existing vessels (by VEGF acting as a survival factor for newly formed vessels²⁴, or owing to deficient or incorrect blood flow).

Several novel findings have emerged from this study. (1) The delayed endothelial cell differentiation in VEGF deficiency is different from the aborted endothelial cell development in Flk-1 deficiency¹², suggesting either the existence of other (unknown) Flk-1 ligand(s), or the rescue of the Flk-1-related defect by maternal VEGF. This is unlikely, however, because of its paracrine mode of action^{10,11} and the abnormal vascular development inside $VEGF^{+/-}$ or $VEGF^{-/-}$ T-ES-cell-derived embryos, despite the presence in the yolk sac of wild-type (tetraploid) endoderm cells^{16,17} that produce VEGF^{10,11}. (2) Heterozygous VEGF deficiency constitutes perhaps the most severe haploid-insufficient (autosomal) phenotype reported to date, which does not appear to be caused by imprinting of the residual wild-type allele (because of the significant expression of VEGF in $VEGF^{+/-}$ embryos), nor by interference with a dominant-negative VEGF mutant peptide (see Supplementary Information). This implicates a ligand dose-dependent effect of VEGF during vascular development, and the production of minimally required levels of VEGF during embryonic development in a highly regulated manner. (3) The similarity between phenotypes of embryos generated by aggregation of heterozygous mutant ES cells with tetraploid hosts, and embryos obtained by germline transmission, demonstrates the validity of this method to determine the phenotype of genetically altered embryos. This technology is fast and relatively inexpensive (as it does not require animal breeding), and separates extraembryonic and embryonic phenotypes by providing wild-type extraembryonic membranes to the mutant embryo^{16,17}. It has also allowed to study the consequences of homozygous deficiency of a gene that is already embryonic lethal in heterozygous-deficient embryos. □

Received 30 November 1995; accepted 29 January 1996.

1. Klagsbrun, M. & Soker, S. *Curr. Biol.* **3**, 699–702 (1993).
2. Dvorak, H. F., Brown, L. F., Detmar, M. & Dvorak, A. M. *Am. J. Path.* **146**, 1029–1039 (1995).
3. Plate, K. H., Breier, G. & Risau, W. *Brain Path.* **4**, 207–218 (1994).
4. Neufeld, G., Tessier, S., Gitay-Goren, H., Cohen, T. & Levi, B.-Z. *Prog. Growth Factor Res.* **5**, 89–97 (1994).

FIG. 3 a, b, Whole-mount LacZ staining revealing a dense network of sprouting vessels in the head mesenchyme and between the somites (arrows indicate intersomitic vessels), and the proper connections between the dorsal aortae and the heart of a 9.5-d.p.c. F_1 $VEGF^{-/-}$: $tie-1-lacZ^{-/-}$ embryo (a). The 9.5-d.p.c. F_1 $VEGF^{+/-}$: $tie-1-lacZ^{+/-}$ embryo (b) contained a dorsal aorta that appeared as a vascular string with a reduced diameter, fewer intersomitic vessels, and abnormal organization of vessels, connecting with the heart. c, d, Injection of ink made visible the heart and the major intra-embryonic vessels (including the dorsal aorta; arrows in c) in the $VEGF^{-/-}$, but only the heart in the $VEGF^{-/-}$ T-ES cell-derived embryo (both at 9.5-d.p.c.). e, f, Macroscopic examination of the yolk sac revealing the presence of a dense capillary plexus with larger collecting vessels (indicated by arrow) in the 9.5-d.p.c. F_1 $VEGF^{+/-}$ yolk sac (e), but only an irregular and less dense network of wider capillaries without larger collecting vessels in the 9.5-d.p.c. F_1 $VEGF^{-/-}$ yolk sac (f). g, Section through a yolk sac, revealing an individual yolk-sac vessel with endothelial cells (arrowhead), and the visceral endoderm layer (arrow). Lines denote the cross-sectional length of the yolk-sac vessel, adjacent and parallel to the endoderm layer (c), the empty space between vessels (e) and the total yolk-sac length analysed (t). Measurements of these distances were used to calculate the vessel density (see Table 1b). h, i, Sections through a 9.5-d.p.c. F_1 - $VEGF^{-/-}$ yolk sac (h), revealing a network of small vessels and a larger collecting vessel (open arrow), that connects with the intra-embryonic vitelline vessel. Sections through a 9.5-d.p.c. F_1 - $VEGF^{+/-}$ yolk sac (i) revealed fewer but enlarged small yolk-sac vessels, and the absence of the larger collecting vessels.



METHODS. T-ES-derived embryos were dissected at 9.5-d.p.c. and filtered ink was immediately injected into the beating heart, using a heat-drawn capillary and a Narashige micro-syringe-based injector. Dissection, fixation and staining of the embryos were done as described in Fig. 2.

Heterozygous embryonic lethality induced by targeted inactivation of the VEGF gene

Napoleone Ferrara, Karen Carver-Moore, Helen Chen, Mary Dowd, Lucy Lu, K. Sue O'Shea*, Lyn Powell-Braxton, Kenneth J. Hillan & Mark W. Moore

Departments of Cardiovascular Research, Molecular Biology and Bioanalytical Technology, Genentech Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

* Department of Anatomy and Cell Biology, University of Michigan, Ann Arbor, Michigan 48109, USA

ANGIOGENESIS is required for a wide variety of physiological and pathological processes¹. The endothelial cell-specific mitogen vascular endothelial growth factor (VEGF)^{2,3} is a major mediator of pathological angiogenesis⁴⁻⁶. Also, the expression of VEGF and its two receptors, Flt-1 and Flk-1/KDR, is related to the formation of blood vessels in mouse and rat embryos⁷⁻¹⁰. Mice homozygous for mutations that inactivate either receptor die *in utero* between days 8.5 and 9.5 (refs 11,12). However, ligand(s) other than VEGF might activate such receptors^{13,14}. To assess the role of VEGF

- Ferrara, N. *Trends cardiovasc. Med.* **3**, 244-250 (1993).
- Shibuya, M. et al. *Oncogene* **5**, 519-524 (1990).
- Devries, C. et al. *Science* **255**, 989-991 (1992).
- Millauer, B. et al. *Cell* **72**, 835-846 (1993).
- Mustonen, T. & Alitalo, K. *J. Cell Biol.* **129**, 895-898 (1995).
- Breier, G., Albrecht, U., Sterner, S. & Risau, W. *Development* **114**, 521-532 (1992).
- Breier, G., Clauss, M. & Risau, W. *Dev. Dynam.* **204**, 228-239 (1995).
- Shalaby, F. et al. *Nature* **376**, 62-66 (1995).
- Fong, G.-H., Rossant, J., Gertsenstein, M. & Breitman, M. L. *Nature* **376**, 66-70 (1995).
- Maglione, D., Guerrero, V., Viglietto, G., Delli-Bovi, P. & Persico, M. G. *Proc. natn. Acad. Sci. U.S.A.* **88**, 9276-9271 (1991).
- Park, J. E., Chen, H. H., Winer, J., Houck, K. A. & Ferrara, N. *J. biol. Chem.* **269**, 25646-25654 (1994).
- Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W. & Roder, J. C. *Proc. natn. Acad. Sci. U.S.A.* **90**, 8424-8428 (1993).
- Nagy, A. & Rossant, J. in *Gene targeting. A Practical Approach* (ed. Joyner, A. L.) 147-180 (IRL, Oxford, 1994).
- Tischer, E. et al. *J. biol. Chem.* **266**, 11947-11954 (1991).
- Claffey, K. P., Senger, D. R. & Spiegelman, B. M. *Biochim. biophys. Acta* **1246**, 1-9 (1995).
- Potgen, A. J. G. et al. *J. biol. Chem.* **269**, 32879-32885 (1994).
- Friedrich, G. & Soriano, P. *Genes Dev.* **5**, 1513-1523 (1991).
- Puri, M., Rossant, J., Alitalo, K., Bernstein, A. & Partanen, J. *EMBO J.* **14**, 5884-5891 (1995).
- Risau, W. & Flamme, I. A. *Rev. Cell. Dev. Biol.* **11**, 73-92 (1995).
- Alon, T. et al. *Nature Med.* **1**, 1024-1028 (1995).
- Tybuiewicz, V. et al. *Cell* **65**, 1153-1163 (1991).
- Carmeliet, P. et al. *J. clin. Invest.* **92**, 2746-2755 (1993).
- Risau, W. et al. *Development* **102**, 471-478 (1988).
- Stavri, G. T. et al. *FEBS Lett.* **358**, 311-315 (1995).

SUPPLEMENTARY INFORMATION. Available at Nature's World-Wide Web site <http://www.nature.com>; requests may also be addressed to Mary Sheehan at the London editorial office of Nature.

ACKNOWLEDGEMENTS. We thank K. Alitalo, A. Bernstein, J. Partanen, M. Puri and J. Rossant for the *tie-1-lacZ* transgenic mice; A. Damert for sequence information of the murine VEGF genomic clone; J. Herz for the JH₁ ES cells; and M. Baes, G. Carmeliet, M. Clauss, E. Conway, C. Lobe, M. Minne, J. Rossant and M. Sass and the other members of the CTTGT for help and discussions.

CORRESPONDENCE and requests for materials should be addressed to D. C. (e-mail: desire.collier@med.kuleuven.ac.be).

directly, we disrupted the VEGF gene in embryonic stem cells. Here we report the unexpected finding that loss of a single VEGF allele is lethal in the mouse embryo between days 11 and 12. Angiogenesis and blood-island formation were impaired, resulting in several developmental anomalies. Furthermore, VEGF-null embryonic stem cells exhibit a dramatically reduced ability to form tumours in nude mice.

We constructed a targeting vector in which the coding sequence of exon 3 of the VEGF¹⁵ gene was replaced by a neomycin-resistance¹⁶ (*neo^r*) gene (Fig. 1a). Two independent embryonic stem (ES) cell clones with a single VEGF allele mutation were identified (Fig. 1b), one of which was used to generate *VEGF*^{-/-} ES cells (Fig. 1c). To confirm inactivation of the VEGF gene, we analysed the production of VEGF in ES cells by a radioreceptor assay¹³, using the two VEGF receptors Flt-1 and Flk-1/KDR. ES cells subjected to a similar selection procedure to inactivate both alleles of the *c-mpl* gene, which encodes the thrombopoietin receptor¹⁷, were included as controls. Wild-type and *c-mpl*^{-/-} ES cells secrete similar amounts of activity able to compete with both receptors (~15 ng ml⁻¹ after 48 h in culture); *VEGF*^{-/-} cells secreted approximately half of that amount, whereas *VEGF*^{-/-} cells released no VEGF. These findings confirm VEGF inactivation at the protein level and also argue against the possibility that the targeted protein may exert a dominant-negative effect on VEGF release in *VEGF*^{+/-} cells.

We tested the hypothesis that VEGF mediates vasculogenesis *in vitro* by inducing embryoid body formation¹⁸ in wild-type or *VEGF*^{-/-} ES cells. Immunostaining with an antibody specific for CD34 revealed a vascular-like pattern in wild-type embryoid bodies, in agreement with previous studies¹⁸, whereas embryoid bodies

derived from *VEGF*^{-/-} ES cells failed to develop any organized CD34-positive vessel-like pattern (unpublished observations).

We then used the nude mouse model to determine the role of VEGF in ES cell angiogenesis *in vivo* and in tumorigenesis¹⁹. Wild-type or *c-mpl*^{-/-} ES cells formed rapidly growing tumours, whereas *VEGF*^{-/-} ES cells had dramatically impaired ability to grow *in vivo*. The difference in tumour weight between *c-mpl*^{-/-} and *VEGF*^{-/-} clones was more than tenfold by four weeks after injection (Fig. 2a). Histologically, all of the ES-cell-derived tumours were germ-cell tumours containing a variety of differentiated and undifferentiated tissues, consistent with a teratoma²⁰. Although blood vessels were identified in all groups, the number of vessels in the *VEGF*^{-/-} group was strikingly reduced and showed a much less complex branching pattern than that observed in groups with at least one intact VEGF allele (Fig. 2b-e).

VEGF^{+/-} ES cells were used to generate chimaeric mice^{21,22}, and both independent clones used transmitted the ES genetic information to the germ line, as assessed by agouti offspring. However, no viable heterozygous mice were obtained after > 100 offspring, suggesting either that the loss of VEGF was lethal in the heterozygous state or that VEGF is maternally imprinted, which would effectively eliminate expression as the ES cells used are male. In blastocysts collected from females bred to germline-producing males, the mutant allele was detected by the polymerase chain reaction (PCR) at a mendelian ratio, confirming that faulty chromosomal transmission or loss in the ES cells was not responsible for the lack of heterozygous births. Presuming embryonic lethality, we collected embryos derived from germline-producing males at days E9.5–12.5. We observed a defect in blood supply of the yolk sac in several embryos at E9.5–11.5 (Fig. 3a). PCR analysis of these yolk sacs indicated that all aberrant embryos were positive for the *neo^r* gene, and that all normal embryos were wild-type. The *VEGF*^{+/-} embryos exhibited several anomalies²³ (Fig. 3). In the cranial region, the branchial arches were poorly developed and unsegmented. The forelimb buds were positioned at their earlier caudal location and were unsegmented. Also, the forebrain region appeared significantly underdeveloped (Fig. 3b, g). In the heart region, the common atrium and primitive ventricle were developmentally delayed²³, the dorsal aortae were rudimentary, and the thickness of the ventricular wall was markedly decreased (Fig. 3b, c, g, h). Unlike the wild-type, in the mutant

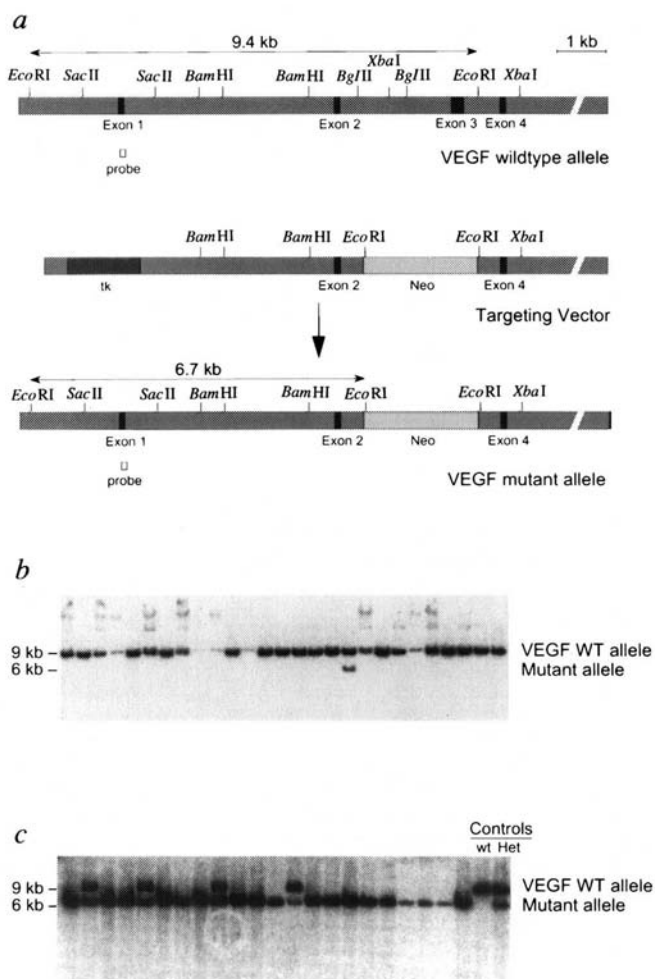


FIG. 1 Targeting vector (a) and Southern blot analysis of single (b) and double (c) allele targeted ES cells.

METHODS. An 18-kb genomic clone encompassing the 5' end of the murine VEGF locus was isolated from a 129 λ FIX library (Stratagene) and characterized according to standard mapping procedures²⁷. A 4.4-kb *Sac*II–*Bgl*II fragment encoding exon 2 and a 7.5-kb *Eco*RI–*Xba*I (vector) fragment containing exon 4 were ligated on either side of a *PGK*–*neo* cassette¹⁶, thus deleting exon 3 (ref. 15). The *Bgl*II site was destroyed in the ligation and an *Eco*RI site was introduced in its place. The direction of transcription of the neomycin-resistance gene was in the opposite orientation to the VEGF gene. A CMV–TKpA cassette was added to the 5' end of the construct to allow negative selection for homologous recombination²⁸. Insertion of the *neo* cassette resulted in the creation of a new *Eco*RI fragment of 6.7 kb in contrast to the 9.4-kb wild-type allele. The targeting vector (20 µg) was linearized and electroporated at 275 V, 200 mF using a BTX 300 electroporator into a subclone of D3 ES cells (ES D3 C12). Cells were subjected to G418 selection¹⁶ at 400 µg ml⁻¹ and 0.2 µM 2-deoxy-2-fluoro-β-D-arabinosyl-5-idouracil (FIAU) for 10 days. Single colonies were expanded for DNA isolation and Southern blot analysis using a 100-bp oligonucleotide probe homologous to exon 1 sequences which are not present in the targeting construct. The targeting frequency was 2 of 1,169 clones screened. The two clones were selected to generate chimaeric mice by microinjection into the blastocoele cavity of 3.5-day C57BL/6J blastocysts^{21,22}. Chimaeric males were mated with C57BL/6J females and agouti offspring were screened for germline transmission by PCR analysis for the *neo^r* gene. No heterozygous mice were obtained. To generate *VEGF*^{-/-} ES cells, a *VEGF*^{+/-} clone was grown in 1.5 mg ml⁻¹ G418 to select for cells that had two copies of the *neo^r* gene²⁹. After 10–12 days, individual clones were expanded for DNA isolation and Southern analysis.

embryo the vitelline veins failed to fuse with the yolk sac (data not shown). The yolk sac revealed a substantially reduced number of nucleated red blood cells within the blood islands (Fig. 3*d, i*). This finding is consistent with there being insufficient activation of the Flk-1/KDR tyrosine kinase by VEGF¹². Haematoxylin and eosin and CD34 immunostaining, revealed significant defects in the vasculature of other tissues and organs including placenta (Fig. 3*e, j*) and the nervous system. The vascular pattern of the forebrain in E10.5 wild-type and heterozygous embryos is shown in Fig. 4*a, b, e, f*. In the wild-type embryos, both the neuroepithelium and the adjacent mesenchyme contain numerous CD34-positive vessels (Fig. 4*a, b*), whereas vascular elements in heterozygotes could be demonstrated in the mesenchyme but not in the neuroepithelium (Fig. 4*e, f*). This failure of blood-vessel ingrowth was accompanied by apoptosis and disorganization of neuroepithelial cells (Fig. 4*e, f*). Consistent with the vascular deficit, apoptosis was dramatically increased in the mutant embryos at E11.5 (data not shown), and death occurs between days 11 and 12. Thus the VEGF^{+/-} embryos survive approximately two days longer than the Flt-1 (ref. 11) or Flk-1/KDR (ref. 12) null embryos, presumably reflecting a partial activation of these tyrosine kinases.

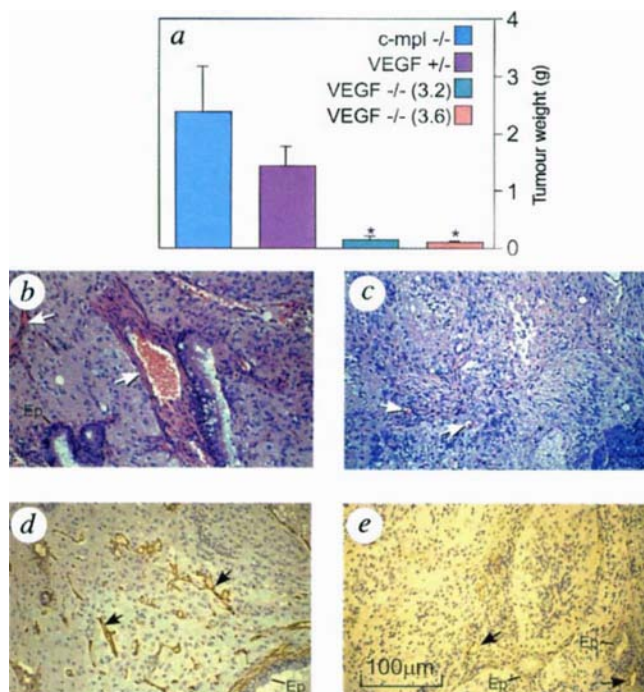


FIG. 2 ES cell tumorigenesis in nude mice. *a*, Tumour weight. *b, c*, Haematoxylin and eosin staining and *d, e*, immunostaining for CD34 of tumour sections derived from *c-mpl*^{-/-} (*b, d*) and VEGF^{-/-} (*c, e*) ES cells. Histology shows representative sections of teratomas, containing glial tissue, blood vessels (arrowheads) and glandular epithelium (Ep).

METHODS. For injection in nude mice, cells that had been cultured in the presence of leukaemia inhibitory factor were dissociated from stock plates by exposure to 125 mM NaCl, 5 mM KCl, 50 mM HEPES, 5 mM glucose, 1 mM EDTA, pH 7.4; they were washed once, and then resuspended in PBS at the appropriate density. Female Beige nude/xid mice (Charles River, Wilmington, DE) 6–10 weeks old were injected subcutaneously with 1×10^7 cells in the dorsal area in a volume of 0.1 ml ($n = 10$). Values shown in *a* are mean $\pm$ s.e.m. of tumour weight determined 4 weeks after cell injection. Statistical analysis was performed by Wilcoxon paired test. Comparison of tumour weight of *c-mpl*^{-/-} group with those of the two VEGF^{-/-} clones revealed P values < 0.005 . The difference in tumour weight between the *c-mpl*^{-/-} and VEGF^{+/-} groups was not significant. For immunohistochemistry, a rabbit polyclonal antiserum raised against murine CD34 was used. Primary antibody was detected with streptavidin–biotin by the Vectastain kit, according to the manufacturer's instructions (Vector).

Specific expression of the VEGF mRNA was detected by *in situ* hybridization in both wild-type and heterozygous embryos. Expression of VEGF mRNA in the neuroepithelium is shown in Fig. 4*c, d, g, h*. Thus the VEGF^{+/-} phenotype appears to be due to gene dosage and not to maternal imprinting. Although several heterozygous phenotypes have been described²⁴, this may be the first report that the loss of a single allele of a gene that does not undergo maternal imprinting can be lethal. It is tempting to speculate that, as VEGF concentrations and angiogenic gradients

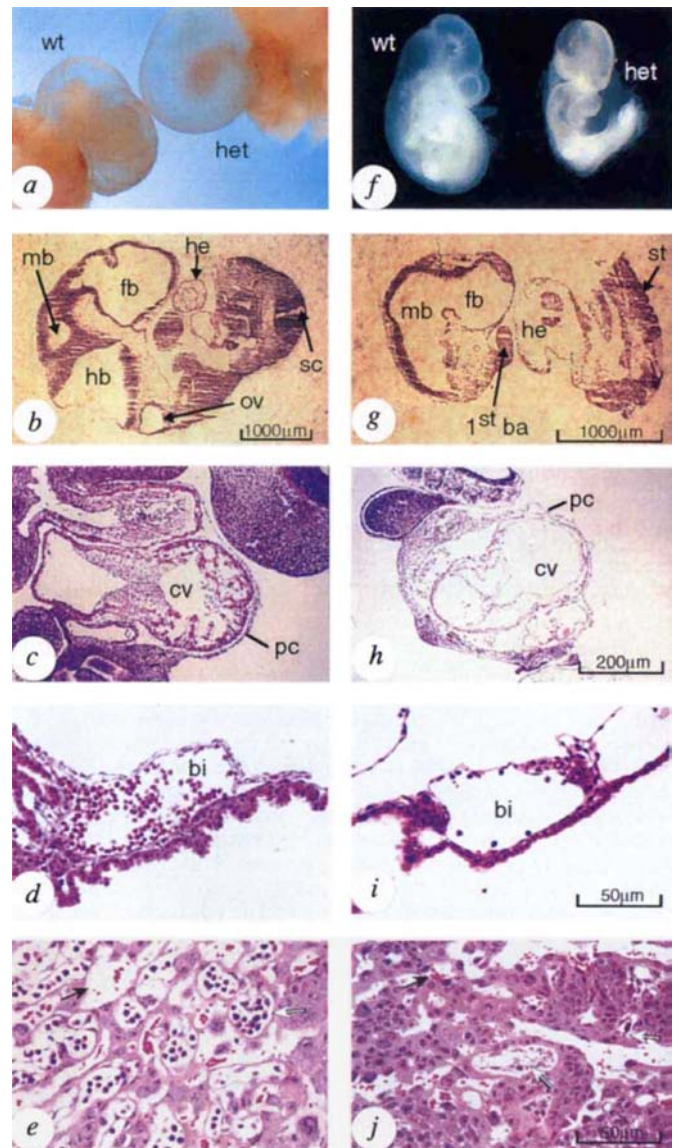


FIG. 3 Gross appearance of wild-type (wt) and heterozygous (het) embryos (*a, f*); haematoxylin and eosin stained sections (*b–e, g–j*). All embryos shown are at E10.5. Note the developmental delay and the apparent absence of blood supply in the yolk sac of the heterozygotes. *b, g*, Sagittal sections of wt (*b*) and het (*g*) embryos illustrating their development. Sections of wt (*c*) and het (*h*) embryos through the cardiac region demonstrate lack of myoblast differentiation in the heterozygote. *d, i*, Yolk sac sections of wt (*d*) and het (*i*) illustrating the scarcity of cellular elements within the blood islands of the het. *e, j*, Sections through the placenta. Note the delicate interweaving of maternal (black arrows) and fetal (white arrows) vessels in the wt placenta (*e*), which contrasts with the condensed appearance and presence of degenerating endothelium lining fetal capillaries in the het (*j*). Abbreviations: hb, hindbrain; mb, midbrain; fb, forebrain; ov, otic vesicle; he, heart; sc, spinal cord; st, somites; ba, branchial arch; pc, pericardium; cv, common ventricle; bi (blood islands).

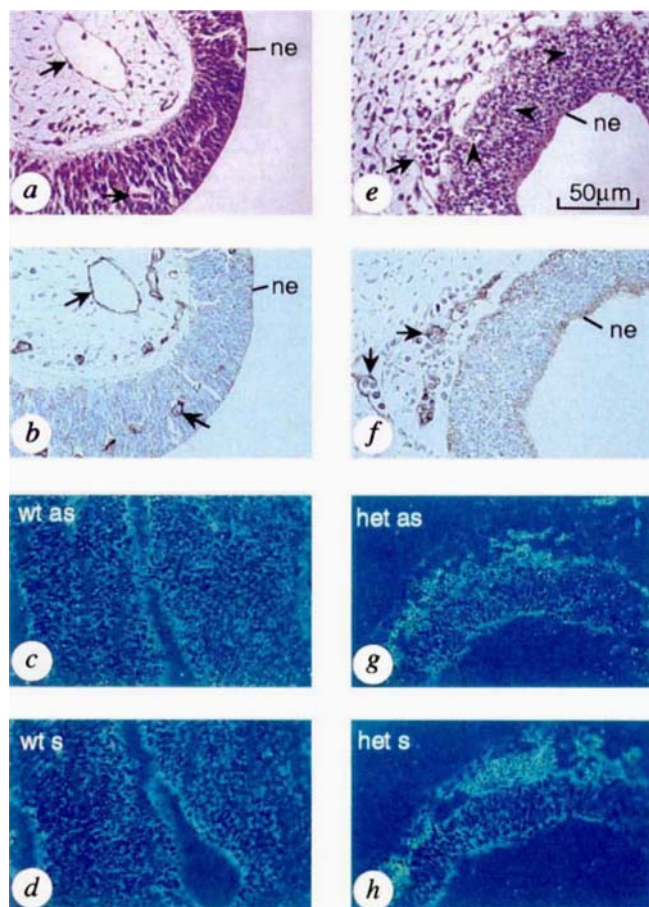


FIG. 4 Haematoxylin and eosin staining (a, e), CD34 immunostaining (b, f) and *in situ* hybridization with a probe specific for VEGF (c, d, g, h) on sections of neuroepithelium (ne) from wild-type (wt) (a–d) and heterozygous (het) (e–h) E10.5 embryos. Arrows indicate blood vessels. Note the absence of blood vessels and the presence of apoptotic cells (arrowheads in e) within the neuroepithelium of the het. This contrasts with the well-differentiated and vascularized neuroepithelium in the wt (a, b). c, d, g, h, Hybridization in both wt and het sections incubated with antisense (as) probe but not in those incubated with control sense (s) probe.

METHODS. CD34 immunostaining was performed on tissues that had been fixed in buffered formalin and then paraffin-embedded as described in Fig. 2. For *in situ* hybridization, PCR amplification was used to generate a 251-bp VEGF fragment with attached T7 and T3 RNA polymerase promoter sequences to allow RNA production. The VEGF probe corresponded to nucleotides 259–509 of the mouse VEGF cDNA⁷. Mouse embryos were fixed in 10% formalin for 2 h and snap-frozen in OCT using hexane, cooled over an acetone/dry ice slurry. Blocks were stored at -80°C , and $\sim 8\text{-}\mu\text{m}$ thick sections were collected on silane-coated slides. Generation of a ^{33}P -UTP labelled antisense and sense RNA hybridization probes and *in situ* hybridization were performed as described³⁰. Slides were developed after two weeks.

fall below a threshold during critical periods, this causes irreversible disruption of normal organogenesis. In this context it is noteworthy that thalidomide, which results in severe limb malformations²⁵ if present even for a short time and at low concentrations in the environment of the early human embryo, has been shown to inhibit angiogenesis²⁶.

The finding that VEGF^{-/-} ES cells are dramatically impaired in their ability to form tumours indicates that, even in a pluripotent system, VEGF is required for effective tumorigenesis and angiogenesis. Finally, although each component of the VEGF/Flt-1/Flk-1 system is required for embryonic development, VEGF is the most critical, as embryonic lethality occurs even in the heterozygous state. □

Received 14 December 1995; accepted 28 February 1996.

1. Folkman, J. & Shing, Y. *J. biol. Chem.* **267**, 10931–10934 (1992).
2. Ferrara, N., Houck, K., Jakeman, L. & Leung, D. W. *Endocr. Rev.* **13**, 18–32 (1992).
3. Dvorak, H. F., Brown, L. F., Detmar, M. & Dvorak, A. M. *Am. J. Path.* **146**, 1029–1039 (1995).
4. Kim, K. J. *et al. Nature* **362**, 841–844 (1993).
5. Adamis, A. P. *et al. Am. J. Ophthalmol.* **118**, 445–450 (1994).
6. Aiello, L. P. *et al. New Engl. J. Med.* **331**, 1480–1487 (1994).
7. Breier, G., Albrecht, U., Sterrer, S. & Risau, W. *Development* **114**, 521–532 (1992).
8. Jakeman, L. B., Armanini, M., Phillips, H. S. & Ferrara, N. *Endocrinology* **133**, 848–859 (1993).
9. Quinn, T., Peters, K. G., deVries, C., Ferrara, N. & Williams, L. T. *Proc. natn. Acad. Sci. U.S.A.* **90**, 7533–7537 (1993).
10. Peters, K. G., deVries, C. & Williams, L. T. *Proc. natn. Acad. Sci. U.S.A.* **90**, 8915–8919 (1993).
11. Fong, G.-H., Rossant, J., Gertenstein, M. & Breitman, M. *Nature* **376**, 66–70 (1995).
12. Shalabi, F. *et al. Nature* **376**, 62–66 (1995).
13. Park, J. E., Chen, H., Winer, J., Houck, K. A. & Ferrara, N. *J. biol. Chem.* **269**, 25646–25654 (1994).
14. Maglione, D., Guerriero, V., Viglietto, G., Delli-Bovi, P. & Persico, M. G. *Proc. natn. Acad. Sci. U.S.A.* **88**, 9267–9271 (1991).
15. Tisher, E. *et al. J. biol. Chem.* **266**, 11947–11954 (1991).
16. Tybulewicz, V. L. J., Crawford, C. E., Jackson, P. K., Bronson, R. T. & Mulligan, R. C. *Cell* **65**, 1153–1163 (1991).
17. Vigon, I. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **89**, 5640–5644 (1992).
18. Risau, W. *et al. Development* **102**, 471–478 (1988).
19. Hilberg, F. & Wagner, E. F. *Oncogene* **7**, 2371–2380 (1992).
20. Pugh, R. B. C. *Pathology of the Testis* (Blackwell, Oxford, 1976).
21. Hogan, B., Beddington, R., Constantini, F. & Lacy, E. *Manipulating the Mouse Embryo* 2nd edn (Cold Spring Harbor Laboratory Press, NY, 1994).
22. Bradley, A. in *Teratocarcinoma and Embryonic Stem Cells: A Practical Approach* (ed. Robertson, E. G.) 113–152 (IRL, Oxford, 1987).
23. Theiler, K. *The House Mouse. Atlas of Embryonic Development* (Springer, New York, 1989).
24. Brandon, E. P., Idzerda, R. L. & McKnight, G. S. *Curr. Biol.* **5**, 625–634 (1995).
25. McBride, W. G. *Lancet* **ii**, 1358–1364 (1987).
26. D'Amato, R. J., Loughnan, M. S., Flynn, E. & Folkman, J. *Proc. natn. Acad. Sci. U.S.A.* **91**, 4082–4085 (1994).
27. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning, a Laboratory Manual* 2nd edn (Cold Spring Harbor Laboratory Press, NY, 1989).
28. Mansour, S. L., Thomas, K. R. & Capecchi, M. R. *Nature* **336**, 348–352 (1988).
29. Mortensen, R. M., Conner, D. A., Chao, S., Geisterfer-Lowrance, A. A. & Seldman, J. G. *Molec. cell. Biol.* **12**, 2391–2395 (1992).
30. Lu, L. H. & Gillett, N. *Cell Vision* **1**, 169–176 (1994).

ACKNOWLEDGEMENTS. We thank L. Berleau and B. Keyt for radioreceptor assays; P. Young and L. Lasky for *in vitro* experiments on embryoid bodies; J. Park for isolation of cDNA probe; R. Taylor and P. Tobin for histology; A. Ryan for comments and advice; D. Giltinan for statistical analysis; and L. Tamayo and D. Wood for graphics.

CORRESPONDENCE and requests for materials should be addressed to N.F.

Convergence of magno- and parvocellular pathways in layer 4B of macaque primary visual cortex

Atomu Sawatari & Edward M. Callaway

Molecular Neurobiology Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA

EARLY visual processing is characterized by two independent parallel pathways: the magnocellular stream, which carries information useful for motion analysis, and the parvocellular stream, which carries information useful for analyses of shape and colour^{1–3}. Although increasing anatomical and physiological evidence indicates some degree of convergence of the two streams, the pathway through layer 4B of primary visual cortex (V1) and on to higher cortical areas is usually considered to carry only magnocellular input^{1–3}. This is inferred from anatomical descriptions of local circuitry in V1, and functional studies of area MT, which receives input from layer 4B^{4–7}. We have directly measured the sources of local functional input to individual layer 4B neurons by combining intracellular recording and biocytin labelling with laser-scanning photostimulation^{8–10}. We found that most layer 4B neurons receive strong input from both magnocellular-stream-recipient layer 4C α neurons and parvocellular-stream-recipient layer 4C β neurons. Thus higher cortical areas that receive input either directly or indirectly from layer 4B are likely to be more strongly influenced by the parvocellular pathway than previously believed.

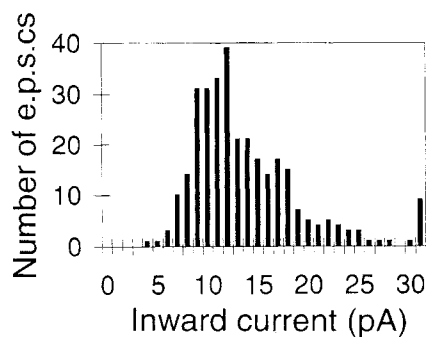


FIG. 1 Histogram showing the size distribution of photostimulation-induced e.p.s.cs measured from a single layer 4B pyramidal neuron. Results are for 312 e.p.s.cs following stimulation at 195 sites. Peak inward currents ranged from 3.4 to 84.5 pA, and averaged 13.6 pA; the median response size was 11.8 pA. Although the mean e.p.s.c. amplitude and smallest detectable responses varied from cell to cell (see text), there was a gradual decrease in response number with current size for every neuron in our sample. Numbers below every fifth bin indicate the upper limit of the peak inward currents for the e.p.s.cs placed in the bin. The last bin includes all e.p.s.cs larger than 30 pA.

METHODS. Brain slices 400 μ m thick were prepared from the primary visual cortex (V1) of macaques. Most of these slices were used for anatomical studies not described here. Slices used in this study were maintained in an interface holding chamber at 34 °C until they were transferred to a recording chamber where they were submerged in ACSF containing 150 μ M 'CNB-caged' glutamate (L-glutamic γ -(α -carboxy-2-nitrobenzyl)ester, trifluoroacetic acid salt)¹⁴ or 250 μ M 'DMNB-caged' glutamate (L-glutamic acid, α -(4,5-dimethoxy-2-nitrobenzyl) ester, hydrochloride) at room temperature. Whole-cell recordings were obtained in layer 4B of V1 brain slices, as described previously¹⁵. The electrodes had resistances of 6–13 M Ω and were filled with the following solution (in mM) 130 potassium gluconate, 1 EGTA, 2 MgCl₂, 0.5 CaCl₂, 2.54 ATP, 10 HEPES, 2% biocytin (Sigma), pH 7.3. After whole-cell recordings were obtained, neurons were voltage-clamped at –65 mV and inward currents evoked by photostimulation were recorded using an Axopatch 1-D amplifier (Axon Instruments). Glutamate was uncaged with 10-ms flashes from an ultraviolet argon-ion laser adjusted for an output of 60–70 mW of ultraviolet light. The laser light was focused through an oil-immersion microscope objective (Nikon, Neo-fluor, 40 $\times$, 1.3 N.A.). The optics were mounted on a computer-controlled motion control system (Klinger) that allowed the scanning of large areas of tissue^{9,10}. The digitized electrical response and the x, y coordinates for each stimulation site were recorded^{9,10}. The peak inward current, the number of e.p.s.cs, and the sum of the peaks of all e.p.s.cs in each trace were measured for the first 250 ms after stimulation.

Extrastriate visual areas are organized into two streams, the parietal stream and the temporal stream. Temporal visual areas receive input (directly and indirectly) from layer 2/3 of the primary visual cortex, V1, and are specialized for the analysis of shape and colour, whereas parietal areas receive input from layer 4B (directly and indirectly) and are specialized for the analysis of motion (see refs 1–3 for reviews). To gain a clearer understanding of the mechanisms of visual processing, it is important to determine how the magnocellular and parvocellular pathways are related to the parietal and temporal streams. In V1, anatomical studies indicate that magnocellular-recipient layer 4C α neurons have strong axonal projections to layer 4B, but parvocellular-recipient neurons in layer 4C β do not^{1–6}. These observations support the long-held notion that layer 4B neurons receive primarily magnocellular input. This interpretation, however, overlooks the possibility that layer 4B neurons could receive input from parvocellular-recipient spiny stellate neurons in layer 4C β onto apical dendrites in layer 3B, or from the unbranched axons in layer 4B itself. To determine whether layer 4B neurons do receive such input, it is necessary to examine functional connectivity; anatomical evidence alone is insufficient.

We studied functional circuitry by measuring excitatory postsynaptic currents (e.p.s.cs) elicited by photostimulation^{8–10} of

neurons in different layers, and combining this with anatomical data about cellular location and morphology. Whole-cell, voltage-clamp recordings (holding potential, –65 mV) were made from neurons in layer 4B of macaque area V1. Slices were bathed in artificial cerebrospinal fluid containing 'caged' glutamate. We uncaged the glutamate locally at desired locations (up to 480 sites spaced 50 μ m apart) with an ultraviolet laser focused to a point through a microscope objective (Fig. 1 legend). The uncaged glutamate elicits action potentials in neurons with cell bodies located near (within 25 μ m or less) the stimulation site (see below). If the stimulated neurons provide synaptic input to the recorded cell, e.p.s.cs are generated. We also iontophoresed biocytin into the recorded neuron to allow later assessment of its morphology and location within the slice (Fig. 2 legend). We obtained photostimulation data for nine layer 4B cells, five of which stained well enough to be identified as either layer 4B pyramidal neurons (four cells) or spiny stellate neurons (one cell). The other four cells were in layer 4B, but because they did not stain well their morphology could not be classified. Here we emphasize responses elicited by stimulation of neurons in layers 4C α and 4C β because they most directly address the question of magnocellular and parvocellular input to layer 4B.

Control experiments indicate that recorded e.p.s.cs are monosynaptic and are driven by neurons near the stimulation site. Figure 2 shows typical voltage-clamp recordings from layer 4B neurons after photostimulation; e.p.s.cs are generated with variable latencies that range from 3 to 250 ms. These responses must all reflect monosynaptic connections from neurons near the stimulation site because spontaneous inward currents were very rare (see below), and action potentials can only be elicited in neurons within 25 μ m of the light flash (see below). Thus secondary action potentials that could give rise to polysynaptic responses do not occur. The variable latency of e.p.s.cs instead results from variability in the time required to initiate action potentials in neurons at the stimulation site; these latencies are similar to those of e.p.s.cs.

Control experiments showed that e.p.s.cs were not due to: (1) random synaptic events that may have occurred simultaneously with the photostimulation; (2) the activation of neurons with cell bodies distant from the stimulation site; or (3) activation of cells independent of the glutamate, that is, activation of neurons by the laser flash. To determine whether random synaptic responses contributed to our measurements, we 'stimulated' cells with the laser off. No e.p.s.cs were observed following stimulation when no laser light was present to uncage the glutamate. To determine whether uncaged glutamate initiated action potentials in neurons distant to the stimulation site, we made current-clamp measurements of electrical responses to direct stimulation. Only when the laser flash was located within 25 μ m of the cell body were action potentials elicited. To determine whether the laser light itself was in some way activating neurons to elicit e.p.s.cs in the recorded cell, we excluded the caged glutamate; the light flashes by themselves did not elicit any detectable responses. These results confirm that the e.p.s.cs observed in the recorded layer 4B neurons were due to activation of neurons at the stimulation sites that made functional connections to the recorded cell, and were not an experimental anomaly.

Synaptic responses, for the most part, were easily detectable over system noise. Figure 1 shows a typical size distribution of e.p.s.cs obtained from a single cell; these data are from the neuron shown in Fig. 24. The histogram combines e.p.s.cs generated by stimulation at all of the sites shown. Most of the responses for all nine cells fell within a distribution spanning 4–30 pA, the size of the smallest detectable response varying from cell to cell (4–7 pA). Although this raises the possibility that we missed some of the smallest responses in some cells, that we saw a gradual decrease in response number with current size makes us confident that we detected all but the smallest and least influential responses.

For all nine cells, we measured the number of e.p.s.cs, the overall peak inward current, and the sum of all e.p.s.cs at their

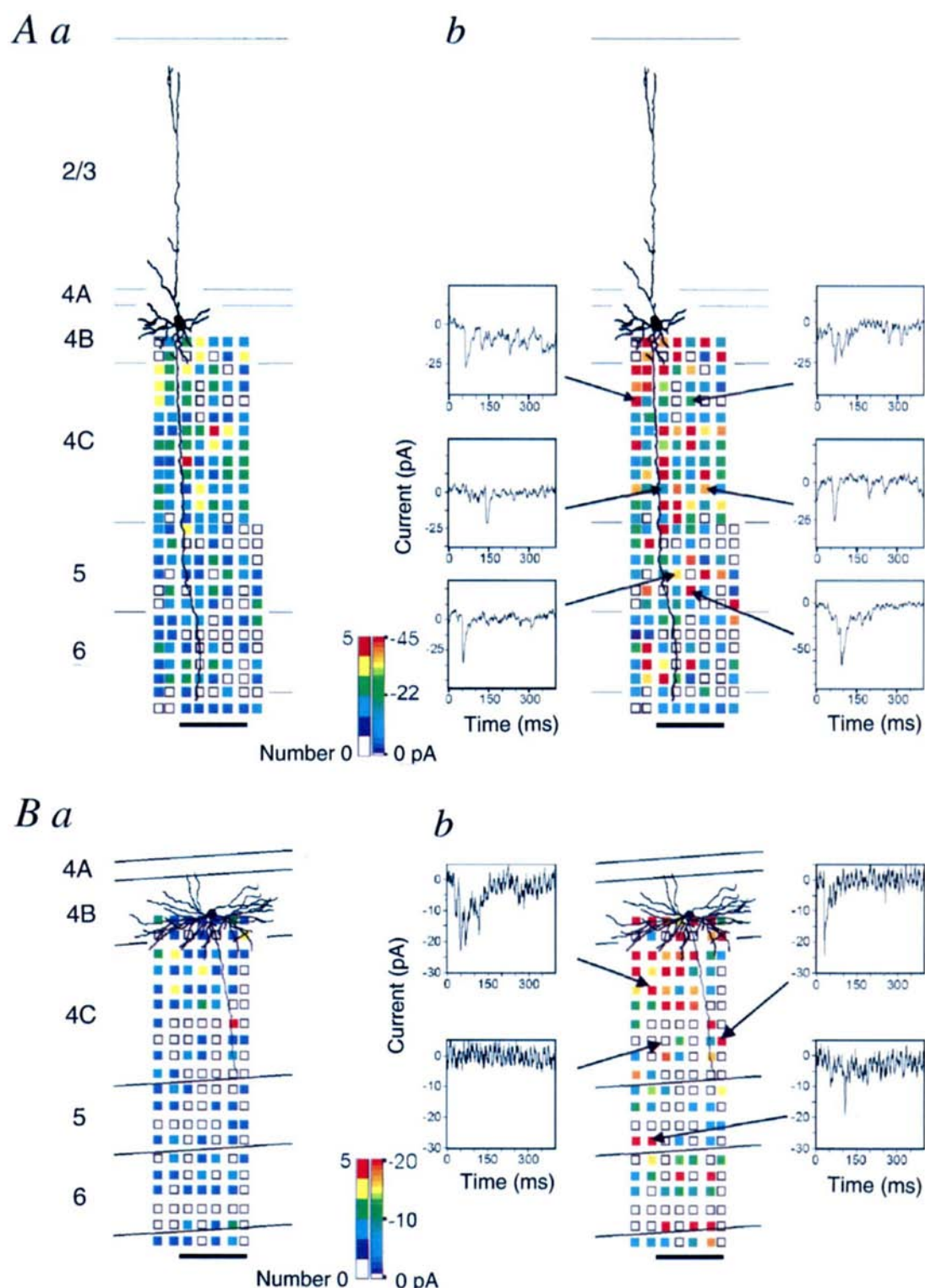


FIG. 2 Distributions of the number (Aa, Ba) and sum of the peaks (Ab, Bb) of e.p.s.cs recorded from a pyramidal (A) and a spiny stellate neuron (B) superimposed on their respective camera lucida drawings. Brief (10 ms) flashes of ultraviolet laser light were focused at each of the locations indicated by pseudo-coloured squares. Such flashes uncage glutamate and evoke action potentials in neurons with somata within 25 μ m of the flash (see text). If any of these neurons make functional connections to the recorded layer 4B cell, one or more e.p.s.cs may be detected. Representative traces illustrating evoked responses are shown in Ab and Bb, and arrows point to the stimulation site corresponding to each trace. We measured

e.p.s.cs of various amplitudes (see Fig. 1). The numbers and amplitudes of e.p.s.cs are indicated by the colours of the squares at each stimulation site, as indicated by the scales at the bottom of A and B; the sites with no detectable responses have white squares. For the pyramidal neuron (A), e.p.s.cs were detected following stimulation at nearly all of the sites in both layers 4C β (lower half of layer 4C) and 4C α (upper half of 4C). These usually included multiple e.p.s.cs (Aa). We did not detect any difference in the numbers of responses evoked following stimulation in 4C α as opposed to 4C β . The amplitude of evoked e.p.s.cs were also similar following stimulation in 4C α and 4C β (Ab). Responses were also detected following stimula-

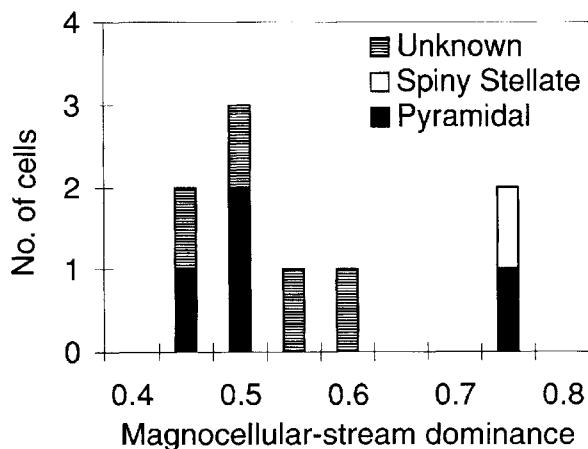


FIG. 3 Histogram showing the distribution of magnocellular-stream dominance for all nine cells. Magnocellular-stream dominance was determined by taking the ratio of the average magnocellular stream ($4C\alpha$) induced response (M) over the sum of the average magnocellular ($4C\alpha$)-(M) and parvocellular-stream ($4C\beta$) (P) induced responses for each recorded cell; $M/(M+P)$. The one spiny stellate cell in our sample, along with one pyramidal cell, exhibited ratios of 0.74 and 0.73, respectively, which correspond to ~ 3 -fold greater magnocellular than parvocellular input. The three remaining identified pyramidal cells, along with four cells of unidentified type, had roughly equal magnocellular and parvocellular input. Filled bars correspond to pyramidal neurons, open bars to spiny stellates, and hatched bars to neurons of unidentified morphological type.

METHODS. The e.p.s.c. amplitudes were averaged over all stimulation sites in layer $4C\alpha$ and then layer $4C\beta$. The average $4C\alpha$ amplitude was then divided by the sum of the average $4C\alpha$ and $4C\beta$ amplitudes to obtain the magnocellular-stream dominance ratio. We averaged over all stimulation sites to reflect both the size and number of inputs from each region.

peaks for each stimulation site. Figure 2 shows the distributions of the number (*Aa*, *Ba*) and sum (*Ab*, *Bb*) of e.p.s.c.s at each stimulation site for a pyramidal cell (*A*) and a spiny stellate cell (*B*) superimposed on camera lucida drawings of their dendritic arborizations and laminar borders. Current traces measured after photostimulation are shown for selected stimulation sites; note that both e.p.s.c. amplitude and latency varied considerably (see above). It was not uncommon to see multiple synaptic events in a single trace (for example, Fig. 2*Ab*, right, middle trace; Fig. 2*Bb*, right, bottom trace). The observed e.p.s.c.s are consistent with the latency and number of action potentials generated by photostimulation in our current-clamp control experiments (see above).

The spiny stellate cell (Fig. 2*B*) has a synaptic input distribution expected for a magnocellular-stream-driven layer 4B neuron. Both the number and sum of responses indicate that the cell receives input predominantly from layer $4C\alpha$. This correlates well with the anatomy, as spiny stellate cells lack an apical dendrite and this, in effect, limits their ability to receive synaptic input from axons projecting to more superficial layers.

Most layer 4B pyramidal cells, however, exhibited an input distribution that was quite unexpected (Fig. 2*A*). Both the sum and number of responses indicate that these cells receive substantial input from both $4C\alpha$ and $4C\beta$. Although the anatomy does not oppose this observation—pyramidal cells do have apical dendrites that branch in layer 3B where neurons in layer $4C\beta$ project their axons—that $4C\beta$ cells contributed such robust synaptic input to the 4B cells was surprising.

To quantify the degree to which $4C\alpha$ and $4C\beta$ cells contributed input to layer 4B neurons, we took the ratio of the average magnocellular ($4C\alpha$) response (average of summed e.p.s.c.s at each stimulation site) over the sum of the average $4C\alpha$ and

average $4C\beta$ responses for each recorded cell. Thus equal inputs correspond to a ratio of 0.5, with larger values being magnocellular-dominated and smaller values parvocellular-dominated. Of the nine cells examined, seven received roughly equal magnocellular and parvocellular input (Fig. 3); these cells had ratios ranging from 0.42 to 0.59. The single spiny stellate cell in our sample, along with one pyramidal cell, received predominantly magnocellular input; their ratios of 0.74 and 0.73, respectively, correspond to approximately threefold greater contributions from $4C\alpha$ than $4C\beta$. Although our small sample size makes it difficult to make any conclusions about the diversity of input ratios, there does seem to be significant parvocellular contribution to cells that had been previously thought to be predominantly driven by the magnocellular stream.

There are potential sources of variability in responses to stimulation at a given site, so it is important to consider how this might influence the ratios we have measured. Such variability could result from variability in the generation of action potentials in neurons at the site, or in the probability of synaptic transmission^{11–13}. Because the ratios we measure incorporate responses averaged over numerous stimulation sites, and both sources of variability are of a stochastic nature, this variability is unlikely to strongly bias the ratio measures. If this is the case, repeated runs through the same stimulation sites in layer 4C should yield similar ratios. We have obtained such measures for a pyramidal neuron, which yielded a ratio of 0.74 during the first trial and a similar value of 0.71 during the second (mean of 0.73; see Fig. 3 and above), and a second neuron of unidentified morphology, which gave a ratio of 0.51 during the first trial and 0.47 during the second (mean of 0.49; see Fig. 3). Thus variability in responses cannot account for the strong parvocellular influence observed in most neurons, nor the stronger magnocellular influence on two of the nine cells.

An updated schematic representation of the local circuitry in macaque V1 is given in Fig. 4. We have demonstrated that there is substantial parvocellular input introduced into the magnocellular stream early in the visual pathways (within V1), and conclude that such a contribution is made to most, but probably not all, layer 4B pyramidal cells. Also, the lone layer 4B spiny stellate neuron in our sample lacks a comparable $4C\beta$ input. However, we cannot exclude the possibility that other spiny stellates, particularly those in more superficial layer 4B, whose dendrites extend into layer 4A, could also receive substantial input from $4C\beta$.

Previous studies have shown that temporal visual areas receive strong input from both magnocellular and parvocellular streams, but feed-forward input to the parietal areas comes from neurons in layer 4B of V1, either directly, or indirectly by means of area V2 thick stripes^{1–3}. Our finding that neurons in layer 4B receive significant input from parvocellular-stream-recipient neurons in $4C\beta$ indicates that parietal areas are probably influenced more by

tion in layers 5 and 6, but these were less common than in layer 4C. For the spiny stellate cell (*B*), responses were detected following stimulation at nearly all sites in layer $4C\alpha$ (24 of 28 sites), but were much less common in $4C\beta$ (11 of 28 sites; *Ba*) and also tended to be smaller (*Bb*). Responses after stimulation of layers 5 and 6 were detected with a similar frequency as those following $4C\beta$ stimulation. Thin black horizontal lines indicate boundaries between layers which are identified by numbers and letters to the left of the figures; Thick black horizontal lines are 250- μ m scale bars.

METHODS. After completion of the physiological analyses, slices were fixed, resectioned, and the sections double-stained for cytochrome oxidase (to reveal laminar borders) and biocytin (to reveal the location and morphology of intracellularly labelled neurons). Dendritic arborizations of labelled cells were reconstructed using a standard camera lucida set-up. Fluorescent microsphere injections were placed in strategic locations within the slice before photostimulation to allow later alignment of the photostimulation data with the anatomical reconstruction.

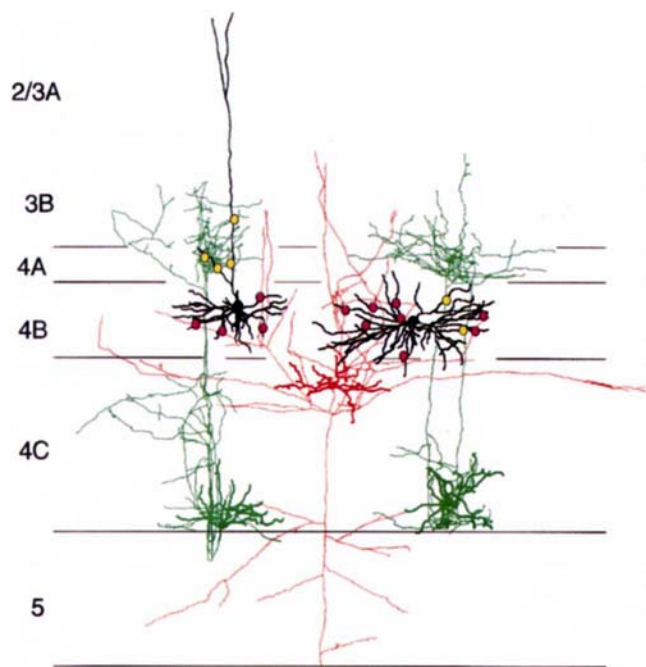


FIG. 4 Schematic drawing showing the organization of local connectivity from layer 4C to individual layer 4B neurons in macaque V1. Layer 4C α spiny stellate neurons (red) have axonal arborizations (finer lines) with branches in layer 4B, whereas layer 4C β spiny stellates (green) lack such branches; instead, the layer 4C β cells project heavily to layers 4A and 3B (ref. 4). Note that layer 4B neurons (black) only dendritic arborizations are shown) with a pyramidal morphology (left) have apical dendritic branches in the vicinity of dense axonal projections from layer 4C β and basal dendritic branches within layer 4B where they can receive input from layer 4C α . The spiny stellate neuron (right) lacks dendritic branches in layers 4A and 3B, and so has minimal overlap with axonal arborizations of layer 4C β neurons, but extensive overlap with projections from 4C α . Coloured spots show the locations of functional synaptic contacts inferred from this study. The pyramidal neuron (left) receives roughly equal numbers of synapses from the magnocellular-dominated layer 4C α (magenta spots) and parvocellular-dominated layer 4C β (yellow spots), but the spiny stellate neuron (right) receives predominantly 4C α input. Black horizontal lines indicate laminar boundaries. Layers are identified by numbers and letters to the left of the figures. This figure is a composite made from reconstructions of intracellularly labelled neurons (unpublished data).

the parvocellular stream than previously believed. This observation does not, however, imply that all parietal areas must be similarly affected. It is important to note that even within our limited sample we found that different cells received different degrees of magnocellular and parvocellular input. Although the influence of parvocellular input to the parietal visual areas has yet to be fully characterized, it seems that the visual system is far more complex than the simple magnocellular and parvocellular parallel pathway model, and that information from both streams contributes importantly to higher-order visual processing in both parietal and temporal visual areas. □

Received 4 October 1995; accepted 26 January 1996.

- Livingstone, M. S. & Hubel, D. H. *Science* **240**, 740–749 (1988).
- Meng, W. H. & Maunsell, J. H. R. *Rev. Neurosci.* **16**, 369–402 (1993).
- Van Essen, D. C. & DeYoe, E. A. *The Cognitive Neurosciences* (ed. Gazzaniga, M. S.) (MIT Press, Cambridge, MA, 1994).
- Lund, J. S. A. *Rev. Neurosci.* **11**, 253–288 (1988).
- Lachica, E. A., Beck, P. D. & Casagrande, V. A. *Proc. natn. Acad. Sci. U.S.A.* **89**, 3566–3570 (1992).
- Yoshioka, T., Levitt, J. B. & Lund, J. S. *Vis. Neurosci.* **11**, 467–489 (1994).
- Maunsell, J. H. R., Nealey, T. A. & DePriest, D. D. *J. Neurosci.* **10**, 3323–3334 (1991).

- Callaway, E. M. & Katz, L. C. *Proc. natn. Acad. Sci. U.S.A.* **90**, 7661–7665 (1993).
- Dalva, M. B. & Katz, L. C. *Science* **265**, 255–258 (1994).
- Katz, L. C. & Dalva, M. B. *J. Neurosci. Meth.* **54**, 205–218 (1994).
- Raastad, M., Storm, J. F. & Andersen, P. *Eur. J. Neurosci.* **4**, 113–117 (1992).
- Allen, C. & Stevens, C. F. *Proc. natn. Acad. Sci. U.S.A.* **91**, 10380–10383 (1994).
- Stevens, C. F. & Wang, Y. *Neuron* **14**, 795–802 (1995).
- Wieboldt, R. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 8752–8756 (1994).
- Blanton, M. G., Lo Turco, J. J. & Kriegstein, A. R. *J. Neurosci. Meth.* **30**, 203–210 (1989).

ACKNOWLEDGEMENTS. We thank L. Katz and M. Dalva for sharing their data acquisition and motion control software, and J. Lieber for technical assistance. This work was supported by the National Eye Institute and the Esther A. and Joseph Klingenstein Fund. A.S. was supported by an NIH training grant and the Chapman Charitable Trust.

Role of intercellular interactions in heterosynaptic long-term depression

Massimo Scanziani*, Robert C. Malenka†‡ & Roger A. Nicoll*†

Departments of * Cellular and Molecular Pharmacology, † Physiology and ‡ Psychiatry, University of California, San Francisco, California 94143, USA

BIDIRECTIONAL control of synaptic strength is thought to be important for the development of neuronal circuits and information storage. The demonstration of homosynaptic long-term depression¹ greatly enhances the usefulness of the synapse as a mnemonic device, but theoreticians have also seen the need for heterosynaptic decreases in synaptic efficacy, both in neuronal development^{2–4} and information storage⁵. Indeed, induction of long-term potentiation in one population of synapses can be associated with a modest depression at neighbouring inactive synapses in the same population of cells^{6–10}. Here we report that in the CA1 region of the hippocampus this heterosynaptic long-term depression has the property that its sites of induction and expression occur in different populations of cells and thus requires the spread of a signal between neurons. Such a mechanism ensures a widespread distribution of this form of plasticity.

Tetanic stimulation of one set of afferents in the CA1 region of the hippocampus routinely elicited a heterosynaptic long-term depression (LTD) in an inactive, neighbouring set of synapses (Fig. 1Aa). This depression averaged $15 \pm 2\%$ ($n = 20$), was non-decremental at 60 min ($n = 5$), and was associated with a long-term potentiation (LTP) of the tetanized pathway (Fig. 1Ab). A depression of similar size could be recorded when the two pathways were on either side of the pyramidal cell layer ($n = 8$). Similar to homosynaptic LTD and LTP¹, heterosynaptic LTD was reversibly blocked by the *N*-methyl-D-aspartate receptor (NMDAR) antagonists (D-(−)-2-amino-5-phosphonopivalic acid (AP5) ($2 \pm 3\%$ depression; $n = 5$; Fig. 1B) and 3-(D)-2-Carboxypiperazin-4-yl-propyl-1-phosphonic acid (CPP) ($20 \mu\text{M}$) ($n = 3$). Heterosynaptic LTD is distinct from transient adenosine-mediated heterosynaptic inhibition^{11,12} in that it is not blocked by adenosine A₁ antagonists (not shown). Furthermore, stopping stimulation of the tetanized pathway ($n = 5$) or depotentiating it with 1 Hz stimulation ($n = 5$) had no effect on the duration or magnitude of heterosynaptic LTD.

One mechanism for affecting synapses at a distance from the activated pathway is the activation of voltage-dependent Ca²⁺ channels by the tetanus-induced depolarization. However, the L-type Ca²⁺-channel blocker nimodipine had no effect on heterosynaptic LTD (Fig. 1C; $14 \pm 2\%$ depression, $n = 7$), contrary to a previous report¹³. This result, nonetheless, does not rule out a role for other classes of Ca²⁺ channel. To eliminate voltage-dependent Ca²⁺ entry during the tetanus, yet avoiding the use of Ca²⁺-blockers that depress synaptic transmission, we held the cell at

positive potentials so that during the tetanus the envelope of synaptic current was actually outward (Fig. 1*Da*). Under these conditions heterosynaptic LTD was undiminished (Fig. 1*D*, filled circles; $33 \pm 5\%$ depression, $n = 6$). In three of these cells, before the tetanus, the depolarization alone had no lasting effects on synaptic strength (Fig. 1*Db*, open circles; $2 \pm 1\%$ depression after 5–8 min).

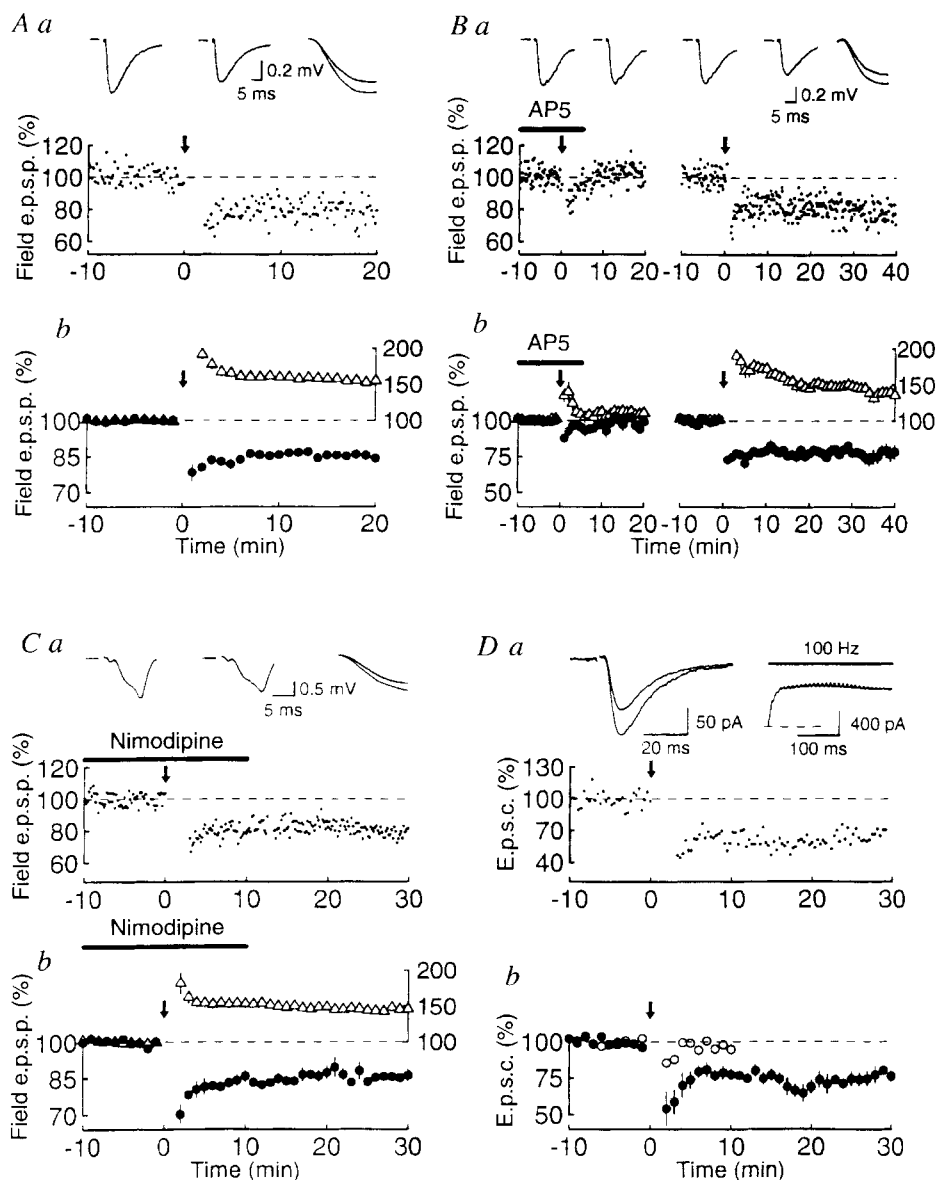
The previous results suggest, to our surprise, that a rise in postsynaptic Ca^{2+} may not be required for heterosynaptic LTD. To test this directly we compared the magnitude of heterosynaptic LTD in cells loaded with the Ca^{2+} chelator 1,2-bis(2-

Aminophenoxy)ethane- N,N,N',N' -tetraacetic acid (BAPTA) (Fig. 2*B*; $n = 8$) to that observed in interleaved control cells (Fig. 2*A*; $n = 5$). Although BAPTA completely blocked LTP (Fig. 2*Bb*, open triangles), heterosynaptic LTD remained, and was actually larger than that recorded in control cells ($31 \pm 3\%$ and $12 \pm 1\%$ depression, respectively). The failure of BAPTA to block heterosynaptic LTD is provocative, because it distinguishes it from homosynaptic LTD¹⁴. We therefore performed a series of experiments documenting that BAPTA did indeed block homosynaptic LTD (Fig. 2*Ca*, top; $4 \pm 6\%$ depression, $n = 4$). This differential action of BAPTA could also be demonstrated in the

FIG. 1 Induction of heterosynaptic LTD requires activation of NMDA receptors but not postsynaptic voltage-gated Ca^{2+} channels. *Aa*, Example of heterosynaptic LTD induced by tetanizing a second, independent pathway at the time indicated by the arrow. The voltage traces represent average field excitatory postsynaptic potentials (f.e.p.s.ps) recorded immediately before and 20 min after the induction protocol. The same traces are shown superimposed, with an expanded time scale, on the right. *b*, Summary graph ($n = 20$) showing both the tetanized pathway (open triangles, right ordinate) and the heterosynaptic pathway (filled circles, left ordinate). *Ba*, Bath perfusion of AP5 (25–50 μM) reversibly blocked the induction of heterosynaptic LTD. *b*, Summary graph ($n = 5$).

Ca, The L-type Ca^{2+} channel blocker nimodipine (50 μM) did not affect the induction of heterosynaptic LTD. *b*, Summary graph ($n = 7$). *Da*, Voltage clamping a cell at +40 mV during the conditioning protocol does not block heterosynaptic LTD. Left, the current traces show two superimposed average e.p.s.cs recorded (at -70 mV) immediately before and 20 min after the conditioning protocol. Right, the average of the four outward currents (recorded at +40 mV) induced by the conditioning tetani. *b*, Similar experiments ($n = 6$) are summarized (filled circles). Depolarization to +40 mV for 90 s without tetanizing did not induce heterosynaptic LTD (open circles).

METHODS. Standard procedures were used to prepare hippocampal slices (400–500 μm) from 2–4-week-old Sprague-Dawley rats²⁷. After dissection, slices were maintained at room temperature for at least 1 h in a submerged chamber containing artificial cerebrospinal fluid (ACSF) equilibrated with 95% O_2 and 5% CO_2 and then transferred to a superfusing chamber for recording. The ACSF contained (in mM) 119 NaCl, 2.5 KCl, 1 NaH_2PO_4 , 1.3 MgSO_4 , 2.5 CaCl_2 , 26 NaHCO_3 , 11 glucose. For whole-cell recordings, the concentration of CaCl_2 and MgSO_4 in the ACSF was raised to 4 mM and 0.5 mM CNQX was added. All experiments were performed at room temperature in the presence of 100 μM picrotoxin. A cut between CA1 and CA3 prevented spread of epileptic activity into CA1. Extracellular fields were recorded with glass electrodes containing 3 M NaCl. Two separate Schaffer collateral inputs were stimulated (100 μs duration) at 0.05–0.10 Hz with two bipolar tungsten electrodes placed in the stratum radiatum on each side of the recording electrode. The amplitude of the field potentials on both pathways ranged from 0.5 to 0.75 mV corresponding to a stimulus strength less than 50% of maximal. Heterosynaptic LTD was induced by a 100 Hz tetanus for 1 s, repeated four times at 20-s intervals. Whole-cell recording electrodes were filled with a solution containing (in mM) 122 Cs-gluconate, 17.5 CsCl,



10 HEPES, 0.2–1 EGTA, 8 NaCl, 2 MgATP, 0.2 cyclic AMP, 0.3 Na_3GTP . In some experiments potassium was substituted for caesium. Data were acquired and analysed as described previously²⁷. All values are expressed as means $\pm$ s.e.m.. The slope and amplitude were measured for field and whole-cell recordings, respectively. Unless stated otherwise, the reported values for heterosynaptic LTD were calculated by taking the average of all responses over a 4-min recording period between 17 and 20 min after conditioning. Drugs used were: picrotoxin, BAPTA, $N\omega$ -nitro-L-arginine (Sigma), CPP, CNQX, AP5, (+)-MCPG (Tocris Cookson), (+)-MK-801, nimodipine (Research Biochemicals), Microcystin-LR (Calbiochem), and FK506 (gift of Fujisawa USA).

same cell (Fig. 2Cb). These experiments establish that, although homosynaptic LTD requires postsynaptic Ca^{2+} , heterosynaptic LTD does not.

The NMDARs that are required to generate heterosynaptic LTD are either located at the synapses expressing heterosynaptic LTD, or are associated with the tetanized pathway. The former would occur if ambient glutamate levels can activate NMDARs while the cell is being depolarized by the tetanus¹⁵. We determined the location of the NMDARs needed for heterosynaptic LTD by taking advantage of the properties of the use-dependent NMDAR antagonist MK-801. In the first set of experiments we gave a series of tetani to one pathway which, in the presence of MK-801, elicited LTP on the first bout of tetani (Fig. 3Aa). Then, 40–50 min after stopping perfusion of MK-801, tetanization of a neighbouring pathway still produced heterosynaptic LTD (Fig. 3Aa, c) in the pathway in which NMDARs had been blocked. Separate experiments using 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX) confirmed that repeated tetanization in MK-801 completely and irreversibly blocked NMDARs (Fig. 3Ab; $n = 3$). Thus NMDARs on the recipient pathway are not required for heterosynaptic LTD.

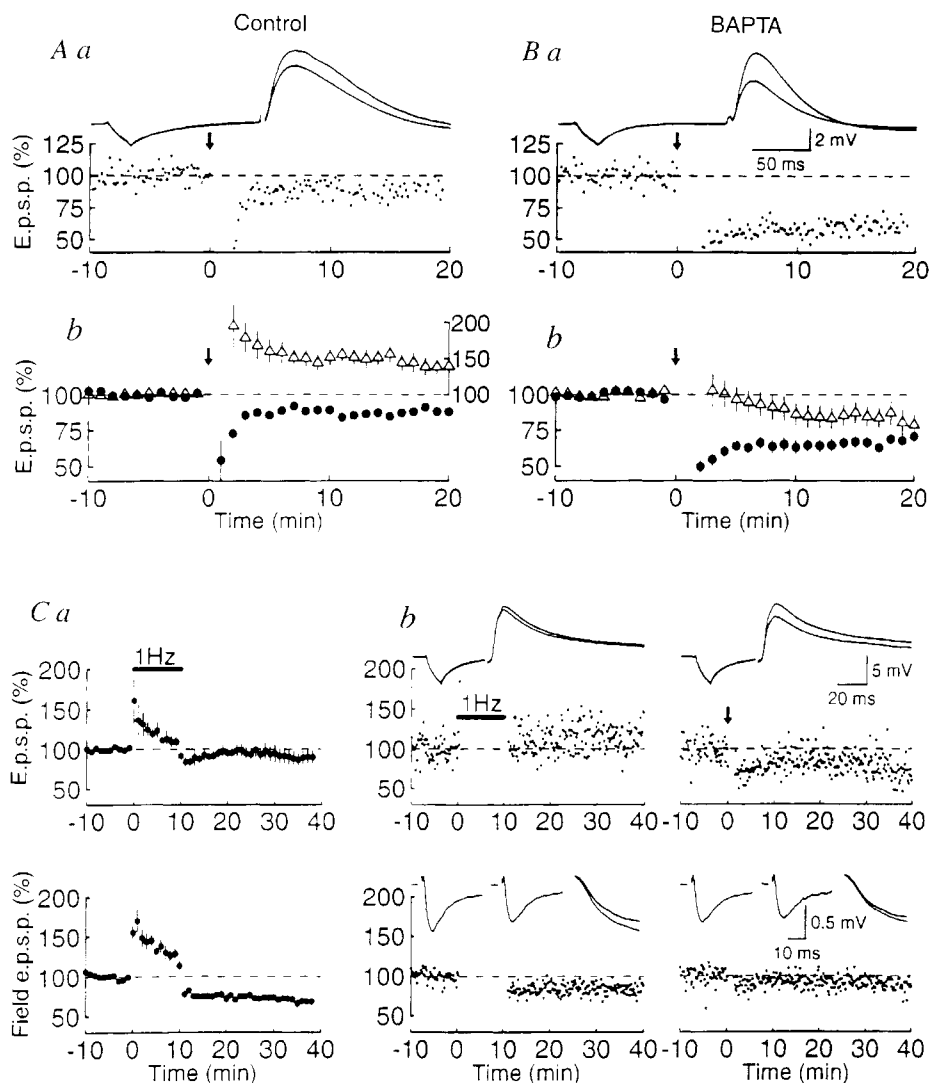
We next examined whether the NMDARs that are required to generate heterosynaptic LTD have to be located on the cell in which the LTD is recorded. We held the cell at positive potentials and activated the input at low frequencies until the synaptic

NMDAR component was blocked (Fig. 3Ba, left). (This procedure does not block NMDARs on neighbouring cells, as these receptors are largely protected from the MK-801 at the resting membrane potential and in the higher concentration of Mg^{2+} used in these recordings.) Tetanizing this pathway still elicited heterosynaptic LTD on a neighbouring input (Fig. 3Ba, right), indicating that normal heterosynaptic LTD was evoked even after complete blockade of NMDARs on the recorded cell ($21 \pm 3\%$ depression, $n = 4$; Fig. 3Bb filled circles). Interestingly, tetanization now evoked a homosynaptic LTD, rather than LTP ($16 \pm 1\%$ depression; Fig. 3Bb, open triangles).

Both nitric oxide (NO)¹⁶ and metabotropic glutamate receptor (GluR) agonists¹⁷ can depress synaptic transmission. However, neither the NO synthase inhibitor *N*- ω -nitro-L-arginine (100 μM ; $n = 4$) nor the metabotropic GluR antagonist (+)- α -methyl-4-carboxyphenylglycine (MCPG) (500 μM ; $n = 5$) blocked heterosynaptic LTD.

Despite their profoundly different induction mechanisms¹⁴, homosynaptic and heterosynaptic LTD seem to converge onto a common process. Thus prior saturation of homosynaptic LTD completely prevents the expression of heterosynaptic LTD (Fig. 4A; $n = 4$). These slices, however, showed heterosynaptic LTD, or more precisely a heterosynaptic depotentiation¹⁸, because subsequent tetanization of the pathway that failed to show heterosynaptic LTD did cause a depression of the pathway that had been

FIG. 2 Intracellular BAPTA does not affect the expression of heterosynaptic LTD but blocks homosynaptic LTD. Aa, Whole-cell current-clamp recording showing heterosynaptic LTD. Superimposed average e.p.s.ps recorded immediately before and 20 min after the conditioning protocol. b, Summary graph ($n = 5$) showing both the tetanized pathway (open triangles, right ordinate) and the heterosynaptic pathway (filled circles, left ordinate). Ba, Experimental conditions as in Aa but with 10 mM BAPTA in the recording electrode. b, Summary graph ($n = 8$). Note the block of LTP in the tetanized pathway (open triangles). Ca, Summary graph of simultaneous field (bottom) and whole-cell recordings with 10 mM BAPTA in the patch electrode (top) ($n = 4$). Stimulation (1 Hz) for 10 min induced homosynaptic LTD of the field e.p.s.p. but not of the e.p.s.p. recorded in the cell. b, Experimental configuration as in a. Intracellular BAPTA blocks homosynaptic LTD (top, left) but not subsequent induction of heterosynaptic LTD (top, right). The field recordings (bottom) show both homosynaptic and heterosynaptic LTD. Membrane potential -70 mV in Aa and Ba, -60 mV in Cb; calibration pulse, -50 pA .



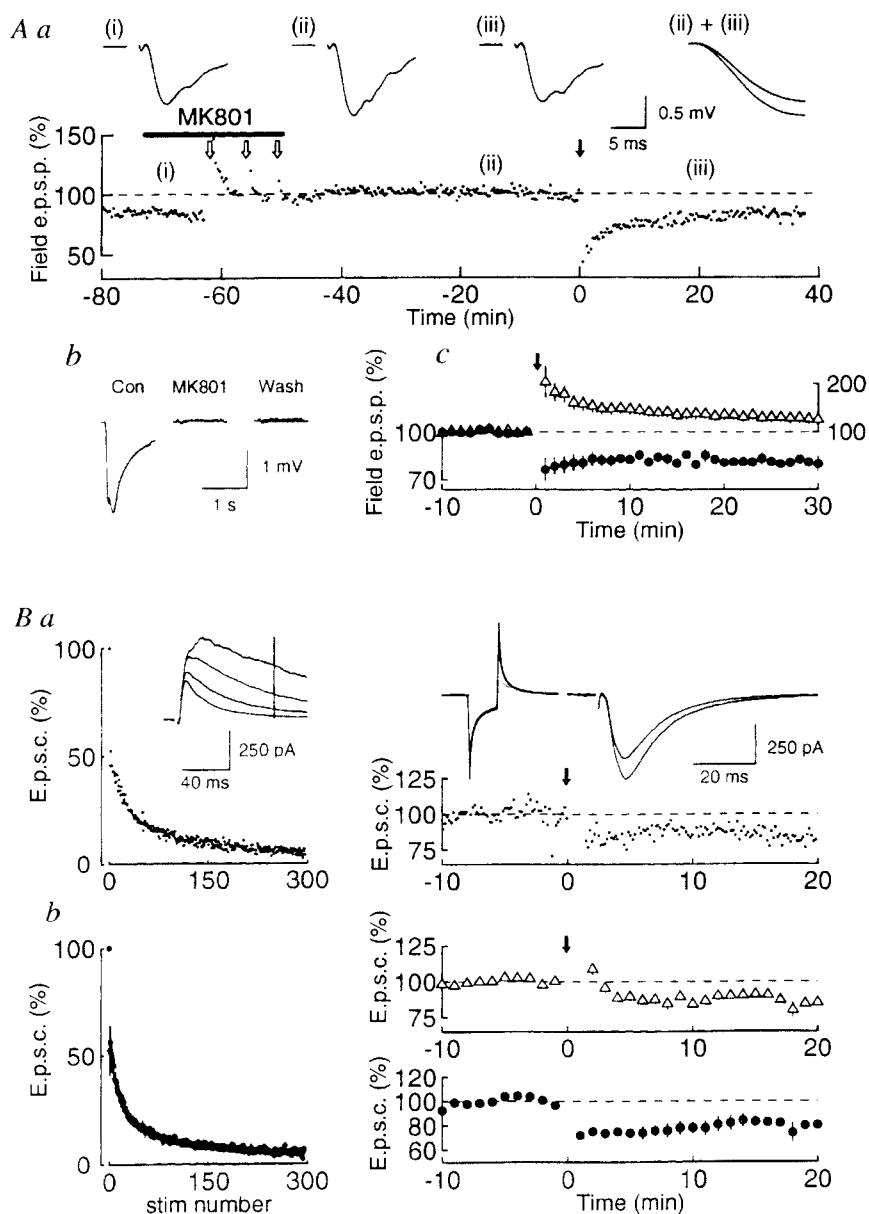


FIG. 3 Heterosynaptic LTD requires intercellular interactions. **Aa**, Repetition of the standard tetanic protocol (see Methods) three times (open arrows) in the presence of MK801 (40 μ M). A subsequent series of four tetani applied to a second independent pathway (filled arrow) was still able to induce heterosynaptic LTD. **b**, NMDAR-mediated field potential evoked with a 100-Hz 1-s stimulation in the presence of CNQX (20 μ M) (Con), after repeated tetanization in the presence of MK801, and 50 min after washout of MK801 (Wash), indicating an irreversible block of NMDARs. **c**, Summary graph showing both the tetanized pathway (open triangles, right ordinate) and the heterosynaptic pathway (filled circles, left ordinate) for a series of experiments as described in **a** ($n = 5$). **Ba**, Whole-cell voltage-clamp recording showing that blockade of NMDARs on the recorded cell does not effect heterosynaptic LTD. **Ba**, Left, the cell was held at +40 mV and one pathway stimulated at 1 Hz for 5 min in the presence of MK801 (40 μ M) to block NMDAR e.p.s.cs. The amplitude of the NMDAR-mediated e.p.s.c. was measured 80 ms after the onset of the response (inset, vertical line). Right, subsequent tetanization of this pathway with the cell voltage clamped at -70 mV was still able to induce heterosynaptic LTD on a second, independent pathway (calibration pulse, -5 mV). **b**, Summary graphs showing complete block of NMDARs in the conditioning pathway (left) and the subsequent ability to generate heterosynaptic LTD in an independent path (right, filled circles) by tetanizing the conditioning path which also now expressed LTD (open triangles) ($n = 4$).

tetanized earlier in the experiment (Fig. 44c, right; $13 \pm 2\%$ depression; $n = 4$).

To further compare the mechanisms of heterosynaptic LTD with those underlying homosynaptic LTD, we examined the effects of two protein phosphatase inhibitors that block homosynaptic LTD. Heterosynaptic, like homosynaptic, LTD was blocked when microcystin-LR, an inhibitor of protein phosphatases 1 (PP1) and 2A (PP2A), was applied intracellularly¹⁹ (Fig. 4B; $5 \pm 6\%$ depression; $n = 7$). In contrast, intracellular application of the calcineurin inhibitor FK506 failed to block heterosynaptic LTD (Fig. 4C; $18 \pm 3\%$ depression; $n = 5$), consistent with the finding that heterosynaptic LTD does not require increases in postsynaptic Ca^{2+} .

These results suggest that tetanic activation of NMDARs on neighbouring cells releases a substance that can activate PP1 in a Ca^{2+} -independent manner. Although little is known about the Ca^{2+} -independent regulation of phosphatases, it has been found that activation of tyrosine kinase by platelet-derived growth factor or epidermal growth factor can stimulate phosphatases in 3T3 cells²⁰. It is important to point out that, while Ca^{2+} channels do not appear to be involved in the heterosynaptic LTD studied here, our

results do not exclude that, under different conditions, Ca^{2+} channels could contribute to heterosynaptic LTD.

The involvement of a diffusible substance ensures that the heterosynaptic LTD is widespread. A diffusible factor may also distribute LTP to neighbouring cells^{21,22}, although in this case the spread appears to be very restricted. That LTP is replaced by LTD after selective blockade of NMDARs on a single cell or buffering postsynaptic Ca^{2+} suggests that heterosynaptic LTD occurs at all synapses, including those expressing LTP. As is the case for NMDAR-dependent LTP²³ and cerebellar LTD²⁴, heterosynaptic LTD was more robust in the absence of inhibition ($15 \pm 2\%$ versus $9 \pm 2\%$, $n = 5$, depression in the presence and absence of picrotoxin, respectively). This is not surprising because electrical stimulation is known to produce stronger inhibition than physiological stimuli²⁵. Moreover, learning processes in rats take place when inhibition in the hippocampus is suppressed²⁶. Heterosynaptic LTD may have been overlooked in previous work because of its relatively small size, especially if factors such as age of the animal and level of disinhibition were not optimized.

The widespread heterosynaptic LTD reported here could initi-

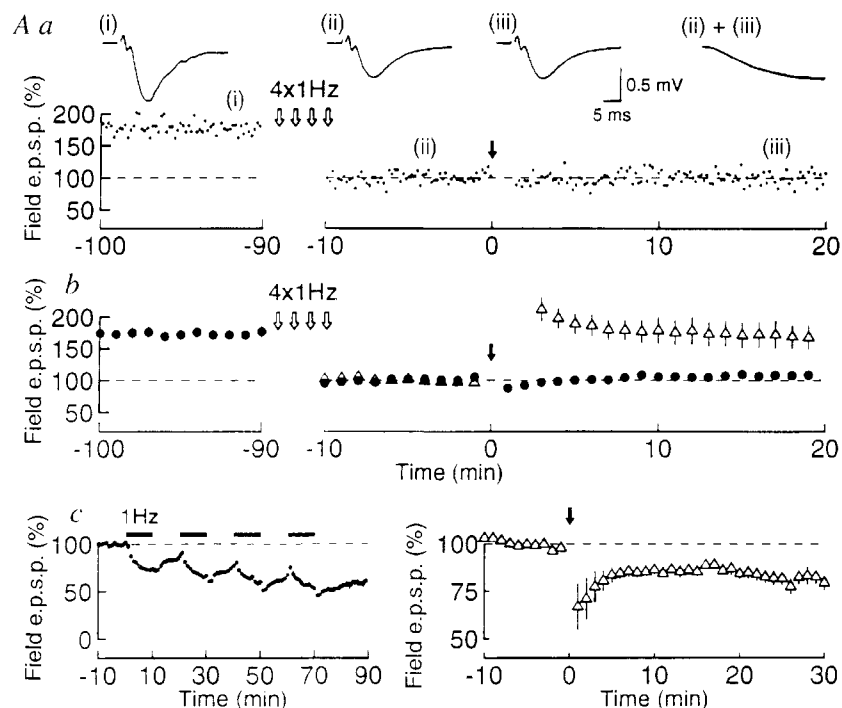
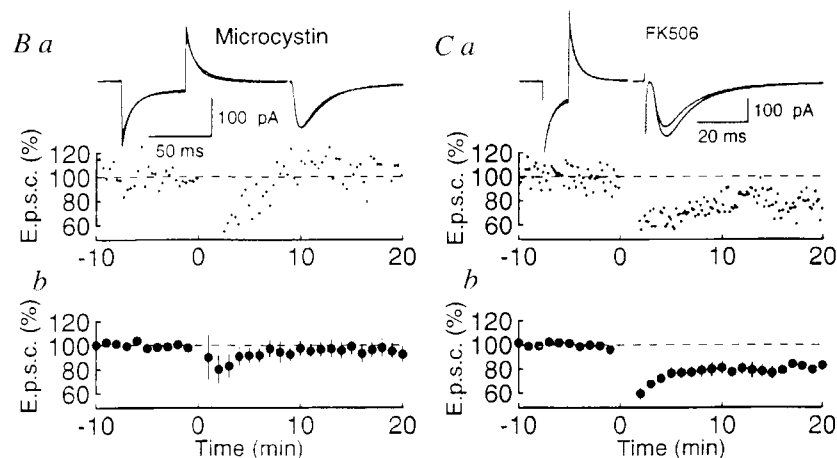


FIG. 4 Heterosynaptic LTD shares some common mechanisms with homosynaptic LTD. *Aa*, A series of 600 stimuli at 1 Hz was repeated four times (open arrows) to saturate homosynaptic LTD. Subsequent tetanization of a second independent pathway (filled arrow) failed to induce heterosynaptic LTD. *b*, Summary graph ($n = 4$). *c*, The induction of homosynaptic LTD for the experiments illustrated in *b*, and, in addition, it shows that heterosynaptic LTD can be induced on the previously potentiated pathway in the same slice (right; same pathway as that plotted with open triangles in *b*). *Ba*, Intracellular microcystin-LR ($10 \mu\text{M}$) blocks heterosynaptic LTD. *b*, Summary graph ($n = 7$). *Ca*, Intracellular FK506 ($50 \mu\text{M}$) does not affect heterosynaptic LTD. *b*, Summary graph ($n = 5$). V_{holding} , -60 mV ; calibration pulse, -5 mV in *Ba* and *Ca*.



ate or contribute to the activity-dependent pruning of excitatory afferents during development, such as that occurring during the formation of ocular dominance columns²⁻⁴. By providing a form of lateral inhibition, it may also serve to sharpen LTP, and thereby enhance the information stored by synapses expressing LTP.

Received 12 October 1995; accepted 29 January 1996.

1. Bear, M. F. & Malenka, R. C. *Curr. Opin. Neurobiol.* **4**, 389–399 (1994).
2. Artoia, A. & Singer, W. *Trends Neurosci.* **16**, 480–487 (1993).
3. Collingridge, G. L. & Singer, W. *Trends pharmac. Sci.* **11**, 290–296 (1990).
4. Stent, G. S. *Proc. natn. Acad. Sci. U.S.A.* **70**, 997–1001 (1973).
5. Sejnowski, T. J. *J. theor. Biol.* **69**, 385–389 (1977).
6. Linden, D. J. & Connor, J. A. *Rev. Neurosci.* **18**, 319–357 (1995).
7. Christie, B. R., Kerr, S. & Abraham, W. C. *Hippocampus* **4**, 127–135 (1994).
8. Lynch, G. S., Dunwiddie, T. & Gribkoff, V. *Nature* **266**, 737–739 (1977).
9. Castro-Alamancos, M. A., Donoghue, J. P. & Connors, B. W. *J. Neurosci.* **15**, 5324–5333 (1995).
10. Desmond, N. L., Colbert, C. M., Zhang, D. X. & Levy, W. B. *Brain Res.* **552**, 93–98 (1991).

11. Grover, L. M. & Teyler, T. J. *Neurosci. Lett.* **154**, 39–42 (1993).
12. Manzoni, O. J., Manabe, T. & Nicoll, R. A. *Science* **265**, 2098–2101 (1994).
13. Wickens, J. R. & Abraham, W. C. *Neurosci. Lett.* **130**, 128–132 (1991).
14. Mulkey, R. M. & Malenka, R. C. *Neuron* **9**, 967–975 (1992).
15. Sah, P., Hestrin, S. & Nicoll, R. A. *Science* **246**, 815–818 (1989).
16. Boulton, C. L. et al. *Eur. J. Neurosci.* **6**, 1528–1535 (1994).
17. Baskys, A. & Malenka, R. C. *J. Physiol., Lond.* **444**, 687–701 (1991).
18. Muller, D., Heft, S. & Figurov, A. *Neuron* **14**, 599–605 (1995).
19. Mulkey, R. M., Herron, C. E. & Malenka, R. C. *Science* **261**, 1051–1055 (1993).
20. Chan, C. P., McNall, S. J., Krebs, E. G. & Fischer, E. H. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6257–6261 (1988).
21. Bonhoeffer, T., Staiger, V. & Aertsen, A. *Proc. natn. Acad. Sci. U.S.A.* **86**, 8113–8117 (1989).
22. Schuman, E. M. & Madison, D. V. *Science* **263**, 532–536 (1994).
23. Wigström, H. & Gustafsson, B. *Nature* **301**, 603–604 (1983).
24. Schreurs, B. G. & Alkon, D. L. *Brain Res.* **631**, 235–240 (1993).
25. Martin, K. A. C. *Q. J. exp. Physiol.* **73**, 637–702 (1988).
26. Wilson, M. A. & McNaughton, B. L. *Science* **261**, 1055–1058 (1993).
27. Castillo, P. E., Weisskopf, M. G. & Nicoll, R. A. *Neuron* **12**, 261–269 (1994).

ACKNOWLEDGEMENTS. We thank P. Salin for input during the course of these experiments, and H. Czerwonka for secretarial assistance. M.S. is supported by the Swiss National Science Foundation and NIMH. R.A.N. is member of the Keck Center for Integrative Neuroscience and the Silvio Conte Center for Neuroscience Research. R.C.M. is a member of the Center for Neurobiology and Psychiatry. R.A.N. and R.C.M. are supported by grants from the NIH.

Direct observation of single kinesin molecules moving along microtubules

Ronald D. Vale^{*†‡}, Takashi Funatsu^{*},
Daniel W. Pierce^{†‡}, Laura Romberg[‡],
Yoshie Harada^{*} & Toshio Yanagida^{*§}

^{*}Yanagida BioMotron Project, ERATO, JRDC, Senba-Higashi 2-4-14, Mino, Osaka 562, Japan

[†]Howard Hughes Medical Institute and [‡]Department of Pharmacology, University of California, San Francisco, California 94143, USA

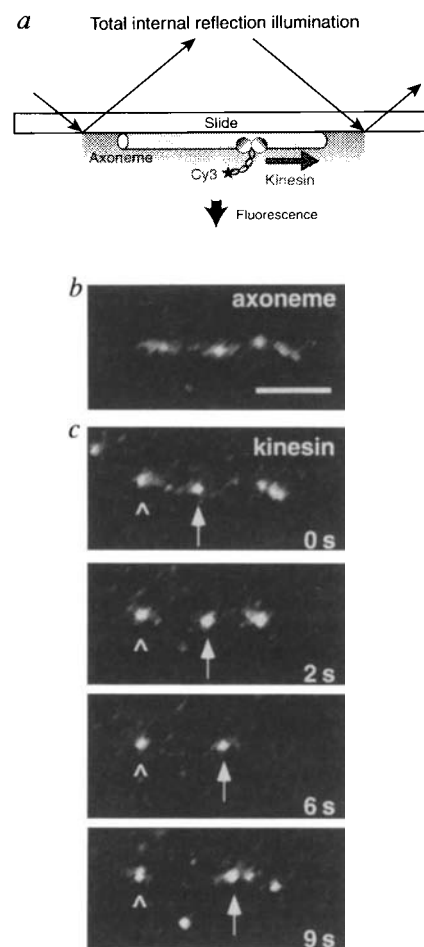
[§]Department of Biophysical Engineering, Osaka University, Toyonaka, Osaka 560, Japan

KINESIN is a two-headed motor protein that powers organelle transport along microtubules¹. Many ATP molecules are hydrolysed by kinesin for each diffusional encounter with the microtubule^{2,3}. Here we report the development of a new assay in which the processive movement of individual fluorescently labelled kinesin molecules along a microtubule can be visualized directly; this observation is achieved by low-background total internal reflection fluorescence microscopy⁴ in the absence of attachment of the motor to a cargo (for example, an organelle or bead). The average distance travelled after a binding encounter with a microtubule is 600 nm, which reflects a $\sim 1\%$ probability of detachment per mechanical cycle. Surprisingly, processive movement could still be observed at salt concentrations as high as 0.3 M NaCl. Truncated kinesin molecules having only a single motor domain do not show detectable processive movement, which is consistent with a model in which kinesin's two force-generating heads operate by a hand-over-hand mechanism.

Kinesin is composed of two heavy chains, each consisting of a 340-amino-acid, N-terminal force-generating domain, a long α -helical coiled coil interrupted in the middle by a proline-glycine-rich hinge, and a small globular C-terminal domain that interacts with light chains^{5,6}. To add a fluorescent label to kinesin in a manner that does not interfere with motor function, the ubiquitous human kinesin gene⁷ was truncated just before the hinge at amino-acid residue 560 and a short peptide containing a reactive cysteine⁸ was introduced at the C terminus (uhK560cys). uhK560cys was expressed in bacteria, purified, and reacted at its C-terminal cysteine with a maleimide-Cy3 fluorescent dye (Fig. 1 legend).

When uhK560-Cy3 (0.6 nM) was combined with 1 mM ATP and Cy5-labelled flagellar axonemes, fluorescent kinesin molecules could be seen binding to and moving unidirectionally along axonemal microtubules (Fig. 1). The velocity of movement was $300 \pm 70 \text{ nm s}^{-1}$ (mean $\pm$ s.d.; $n = 60$), which was similar to the gliding velocity of axonemes on uhK560-Cy3-coated glass surfaces ($370 \pm 40 \text{ nm s}^{-1}$; $n = 28$).

To verify that the translocating objects were single kinesin molecules, the photobleaching characteristics and intensities of fluorescent spots were examined. The fluorescence of surface-adsorbed kinesin disappeared primarily in either a one-step (Fig. 2a) or a two-step process (Fig. 2b), as expected for photobleaching reactions of one or two dye molecules, respectively, bound to the



Cy5-axoneme. c, uhK560-Cy3. The caret (^) marks a stationary fluorescent spot, and the arrows track a moving fluorescence spot (time in seconds). Other spots visible in individual panels are ones that associated briefly ($< 2 \text{ s}$) with the axoneme. Bar, $5 \mu\text{m}$.

METHODS. uhK560 with an additional C-terminal peptide containing a reactive cysteine (PSIVHRKCF)⁸ was expressed in *Escherichia coli*²¹ and purified by phosphocellulose followed by mono-Q chromatography. uhK560-Cy3 (concentration determined from absorbance at 280 nm, A_{280} ; ref. 22) was labelled to a stoichiometry of 0.1 to 0.5 dyes per polypeptide chain as described⁴. The uhK560-Cy3 preparation was centrifuged at 100,000 r.p.m. for 10 min in a table-top ultracentrifuge to remove any aggregates; additional microtubule-affinity purification¹⁹ was sometimes performed to improve motility. Analysis of low concentrations of uhK560-Cy3 by sucrose gradient velocity sedimentation¹⁶ (5 S) and Superose-6 gel filtration chromatography indicates that it is dimeric. The concentrations used in the sucrose gradient (5 nM) suggest that uhK560-Cy3 remains dimerized under the conditions of the assay. The microtubule-stimulated ATPase²³ properties of uhK560 labelled with Cy3 at a 0.43 molar ratio of dye to polypeptide ($k_{\text{cat}} = 22.2 \pm 1.0 \text{ s}^{-1}/\text{site}$; $K_{0.5, \text{MT}} = 1.1 \pm 0.2 \mu\text{M}$) and at 0.96 stoichiometry ($k_{\text{cat}} = 19.2 \pm 0.2 \text{ s}^{-1}/\text{site}$; $K_{0.5, \text{MT}} = 1.6 \pm 0.06 \mu\text{M}$) were similar to that of the unlabelled protein ($k_{\text{cat}} = 25.8 \pm 0.8 \text{ s}^{-1}/\text{site}$; $K_{0.5, \text{MT}} = 0.7 \pm 0.08 \mu\text{M}$) (measured in the motility buffer: 12 mM PIPES, pH 6.8, 2 mM MgCl_2 , 1 mM EGTA). Cy5-labelled axonemes²⁴ were combined with 0.5–1.5 nM uhK560-Cy3 in motility buffer containing 1 mM ATP, 7.5 mg ml^{-1} BSA, 0.5% mercapto-ethanol and an oxygen scavenger system²⁰ at 25°C . BSA inhibited non-specific adsorption of kinesin but did not prevent axonemes from attaching to the slide. Details of low-background total internal reflection fluorescence microscopy are described elsewhere⁴. In these experiments, 5.8 ± 1.0 association events per min per μm axoneme ($n = 8$ axonemes; 5-min observation for each axoneme) were observed when the uhK560-Cy3 concentration was 0.6 nM, whereas only 0.05 ± 0.01 events per min were scored on glass without an axoneme. Kinesin spots moving on an axoneme disappeared at $0.42 \pm 0.03 \text{ s}^{-1}$, which was significantly faster than the rate of photobleaching of glass-attached kinesin ($0.02 \pm 0.0005 \text{ s}^{-1}$).

FIG. 1 Movement of a single fluorescently labelled kinesin molecule along an axoneme. a, Kinesin (uhK560cys) labelled at the C-terminal cysteine residue with Cy3 fluorescent dye was combined with Cy5-labelled axonemes in the presence of ATP. The specimen was illuminated by the evanescent field resulting from total internal reflection of a laser beam at the boundary between the fused silica slide and the aqueous solution⁴. b,

uhK560cys dimer. The percentage of one-step, two-step, three-step and more-than-three-step photobleaching events was 74, 25, 2 and 0%, respectively ($n = 355$). Assuming labelling of only the C-terminal cysteine and random dimerization of two chains, these values agree with the percentages expected of single-dye (72.6%) and double-dye (27.4%) labelled kinesin molecules predicted from the labelling stoichiometry (0.43 dye/polypeptide). The fluorescence intensities of kinesin molecules randomly adsorbed onto a glass slide (Fig. 2c) were also similar to those of moving spots (Fig. 2d). Collectively, these results are consistent with the idea that single kinesin molecules can move along microtubules.

The distance travelled by single kinesin molecules was distributed exponentially (Fig. 3a), as described for bead motility assays⁹. Based upon the mean travel distance of 630 ± 30 nm and an 8-nm step per cycle for kinesin¹⁰, the probability of remaining bound to the microtubule per mechanical cycle is $\sim 99\%$ ($e^{(-8/630)}$). If one ATP is hydrolysed per 8-nm step (which is consistent with uhK560 ATPase rates), then 40 ATP molecules on average would be hydrolysed by each kinesin motor domain per diffusional encounter with the microtubule, which is intermediate between two previous estimates^{2,3}.

The effect of ionic strength on single kinesin molecule motility is shown in Fig. 3. Although the number of measurable associations events decreased dramatically with increasing salt concentration, the average travel distance and velocity (Fig. 3 legend) of kinesin molecules that had initiated movement were much less

TABLE 1 Motility of monomeric and dimeric kinesin

Microtubule gliding		Single molecule fluorescence		
Kinesin	Velocity (nm s ⁻¹)	Velocity (nm s ⁻¹)	Travel distance (nm)	Association events (events per min per μ m MT)
dK340	7 ± 2	—	—	0.08 ± 0.02
dK350	120 ± 10	—	—	0.22 ± 0.07
dK365	110 ± 10	—	—	0.30 ± 0.04
dK440	570 ± 70	340 ± 90	400 ± 40	2.48 ± 0.40

Microtubule translocation induced by monomeric and dimeric *Drosophila* kinesin proteins (number of amino acids indicated) in multiple motor gliding or single fluorescent molecule motility assays. Velocities are the means and s.d. of > 20 measurements (microtubules in the case of the gliding assay, and axonemes in the case of the single molecule fluorescence assay), and the association events are the means and s.d. from > 3 axonemes (5–20 min of observation per axoneme (a minimum of 10 events were scored)). The dK440 travel distance and error were determined as shown in Fig. 3. Dashes signify that no movement was detected. dK340cys, dK350cys, dK365cys, and dK440cys were generated by polymerase chain reaction using the *Drosophila* kinesin gene, expressed in bacteria and purified by phosphocellulose followed by mono-Q chromatography. Hydrodynamic studies indicated that, consistent with previous work^{12,14,15}, dK440 is a dimer and that dK365, dK350 and dK340 are monomers. Sucrose gradient velocity sedimentation indicates that dK440-Cy3 (5S) remains a dimer at 3 nM. For microtubule gliding assays, the kinesin reactive cysteine was modified with long-chain biotin-maleimide⁸, streptavidin was added in an 8:1 ratio, and the kinesin-biotin-streptavidin complexes were isolated by microtubule affinity^{18,19}. For the motility assay, 0.5 mg ml⁻¹ biotinylated BSA was introduced into a microscopic flow chamber. After washing, streptavidin-modified kinesin motor (1–5 μ M) was introduced in motility buffer with 3 mg ml⁻¹ casein. Next, rhodamine-labelled microtubules¹⁸ ($\sim 10 \mu$ g ml⁻¹) were introduced in the above buffer with 1 mM ATP and an oxygen scavenger system²⁰ and examined by fluorescence microscopy. For the single molecule fluorescence assay, *Drosophila* kinesin proteins were labelled with Cy3 and assayed for motility as described for Fig. 1.

affected. The number of ATP molecules hydrolysed per microtubule encounter, however, was reported to decrease 10-fold when salt was raised from 0.05 to 0.15 M (ref. 3). This apparent discrepancy may reflect a difference between bulk ATPase measurement and observation of single molecule behaviour in a

FIG. 2 Photobleaching properties and fluorescence intensities of fluorescently labelled kinesin. uhK560-Cy3 was adsorbed onto a fused-silica slide, and the intensities of individual fluorescent spots plotted against time. Typical examples of one-step (a) or two-step (b) photobleaching events are shown, reflecting one or two Cy3 dye molecules attached to kinesin, respectively. The distribution of the fluorescence intensities of isolated uhK560-Cy3 spots adsorbed onto the slide (c) or moving along the axoneme (d) are comparable, which suggests that moving fluorescent spots are single molecules and not higher-order aggregates. Intensities were measured using a rolling 8-frame average to reduce fluctuations, summing intensities from a 7×7 pixel window, and subtracting the background intensity from an adjacent region on the slide. The intensities of two dye molecules (b) and three dye molecules (not shown) are in the linear range of the camera, as photobleaching caused approximately equal decreases in intensity.

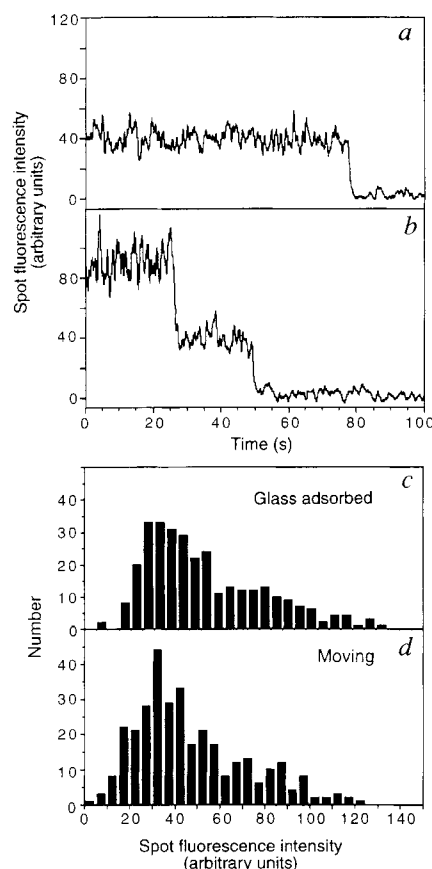
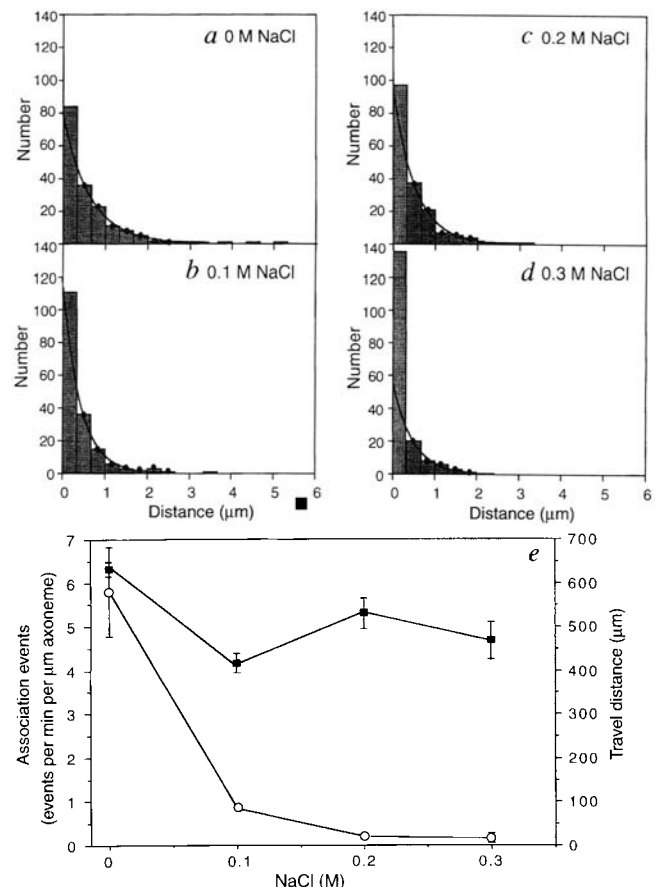


FIG. 3 Effects of ionic strength on the travel distance and association frequencies of uhK560-Cy3. The distances travelled by single uhK560-Cy3 molecules in motility buffer containing 0 M (a), 0.1 M (b), 0.2 M (c) or 0.3 M NaCl (d) are shown. The dots indicate the bins that were used for exponential fitting. In e, the travel distances (filled squares) in a–d are compared with that of uhK560-Cy3 association events with the axoneme (open circles). Mean and standard deviations of association events were scored from 5–12 different axonemes (5–40 associations per axoneme). The velocities of uhK560-Cy3 that travelled for $> 0.5 \mu\text{m}$ were 300 ± 60 , 380 ± 120 , 310 ± 80 and $250 \pm 60 \text{ nm s}^{-1}$ at 0, 0.1, 0.2 and 0.3 M NaCl, respectively ($n = 60, 41, 52$ and 20).

METHODS. Measurements of solitary uhK560-Cy3 molecules were made by tracking fluorescent spots that were positioned over the axoneme for at least 5 video frames. Static associations of $> 3 \text{ s}$ were rare, and were excluded as they probably represented nonspecific associations. Although travel distances of $< 0.3 \mu\text{m}$ could not be measured accurately, all measurable associations, including those with $< 0.3 \mu\text{m}$ translocations, are included in these histograms. As the first bin may contain nonspecific associations with the axoneme, it was excluded from the two-parameter exponential fits. The means determined from the decay constants (630, 415, 520 and 470 nm at 0, 0.1, 0.2, 0.3 M NaCl, respectively) were in general agreement with the distribution averages (620, 430, 500 and 300 nm).



microscopic assay. At higher salt, kinesin may undergo transient microtubule interactions (too fast to be detected in our assay) that result in ATP hydrolysis or product release, but not processive movement. Once kinesin is bound in a manner that allows processive movement to begin, then the travel distance appears to be only modestly affected by ionic strength.

The mechanism by which kinesin moves for long distances on a microtubule without dissociating remains unresolved, although it has been suggested that the two heads of kinesin may operate in a hand-over-hand manner^{3,11–13}. To explore this idea, we prepared a series of truncated, monomeric *Drosophila* kinesin proteins^{12,14,15}, as well as a longer kinesin molecule (dK440) that forms a dimer^{15,16} (Table 1). Consistent with previous studies^{12,17}, even the monomeric motors generated movement in a gliding assay in which many motors interact simultaneously with the microtubule, although the velocity of dK340 was very slow. In the single-molecule fluorescence motility assay, dK440-Cy3 molecules moved along microtubules with a velocity and travel distance comparable to that described for uhK560-Cy3. In contrast to observations for dK440-Cy3, unidirectional movement by

dK340-Cy3, dK350-Cy3 or dK365-Cy3 was not observed. Furthermore, association–dissociation reactions of these monomeric kinesins with the axoneme were seen less frequently than for dK440, suggesting that their association–dissociation cycle may be faster than can be detected by video analysis.

In conclusion, these results provide definitive evidence for the ability of a single kinesin molecule at physiological ionic strength to move along a microtubule even in the absence of attachment to a cargo. The travel distance reported here is only slightly lower than that of kinesin moving a bead of diameter $0.2 \mu\text{m}$ ($1.4 \mu\text{m}$)⁹, which probably reflects the smaller diffusion coefficient of a bead compared with a protein. Furthermore, kinesin molecules containing a single motor domain showed no detectable processive movement, although higher spatial resolution techniques¹⁰ will determine whether they undergo very short movements. In any event, these results and others¹² argue that two kinesin force-generating heads are required for long-distance movements along a microtubule. Assays similar to those described here could be applied to other processive enzymes such as polymerases and helicases. □

Received 24 November 1995; accepted 12 February 1996.

- Bloom, G. & Endow, S. *Motor Proteins 1: Kinesin* (Academic, London, 1994).
- Hackney, D. D. *Nature* **377**, 448–450 (1995).
- Gilbert, S. P., Webb, M. R., Brune, M. & Johnson, K. A. *Nature* **373**, 671–676 (1995).
- Funatsu, T., Harada, Y., Tokunaga, M., Saito, K. & Yanagida, T. *Nature* **374**, 555–559 (1995).
- Yang, J. T., Laymon, R. A. & Goldstein, L. S. *Cell* **56**, 879–889 (1989).
- Hirokawa, N. et al. *Cell* **56**, 867–878 (1989).
- Navone, F. et al. *J. Cell Biol.* **117**, 1263–1275 (1992).
- Itakura, S. et al. *Biochem. biophys. Res. Commun.* **196**, 1504–1510 (1993).
- Block, S. M., Goldstein, L. S. & Schnapp, B. J. *Nature* **348**, 348–352 (1990).
- Svoboda, K., Schmidt, C. F., Schnapp, B. J. & Block, S. M. *Nature* **365**, 721–727 (1993).
- Hackney, D. D. *Proc. natn. Acad. Sci. U.S.A.* **91**, 6865–6869 (1994).
- Berliner, E., Young, E. C., Anderson, K., Mahtani, H. & Gelles, J. *Nature* **373**, 718–721 (1995).
- Schnapp, B. J., Crise, B., Sheetz, M. P., Reese, T. S. & Kahn, S. *Proc. natn. Acad. Sci. U.S.A.* **87**, 10053–10057 (1990).
- Huang, T. & Hackney, D. D. *J. biol. Chem.* **269**, 16493–16501 (1994).
- Correia, J. J., Gilbert, S. P., Moyer, M. L. & Johnson, K. A. *Biochemistry* **34**, 4898–4907 (1995).

- Huang, T., Suhan, J. & Hackney, D. D. *J. biol. Chem.* **269**, 16507–16507 (1994).
- Stewart, R. J., Thaler, J. P. & Goldstein, L. S. *Proc. natn. Acad. Sci. U.S.A.* **90**, 5209–5213 (1993).
- Hyman, A. et al. *Meth. Enzym.* **196**, 303–319 (1990).
- Vale, R. D., Reese, T. S. & Sheetz, M. P. *Cell* **41**, 39–50 (1985).
- Harada, Y., Sakurada, K., Aoki, T., Thomas, D. D. & Yanagida, T. *J. molec. Biol.* **216**, 49–68 (1990).
- Shimizu, T. et al. *Biochemistry* **34**, 13259–13266 (1995).
- Gilbert, S. P. & Johnson, K. A. *Biochemistry* **32**, 4677–4684 (1993).
- Vale, R. D., Coppin, C. M., Malik, F., Kull, F. J. & Milligan, R. A. *J. biol. Chem.* **269**, 23769–23775 (1994).
- Kamiya, R., Kurimoto, E. & Muto, E. *J. Cell Biol.* **112**, 441–447 (1991).

ACKNOWLEDGEMENTS. This work was supported, in part, by grants from the NIH. We thank C. McGee for discussion, T. Yagi for preparation of axonemes, N. Horn-Booher for technical assistance and C. Coppin and R. Cooke for comments on the manuscript.

CORRESPONDENCE and requests for materials should be addressed to R.D.V. (e-mail address vale@phy.ucsf.edu).

RNA editing of hepatitis delta virus antigenome by dsRNA-adenosine deaminase

Andrew G. Polson, Brenda L. Bass & John L. Casey*

Department of Biochemistry and Howard Hughes Medical Institute, 6110a EHG, Building 533, University of Utah, Salt Lake City, Utah 84112, USA

* Georgetown University Medical Center, Division of Molecular Virology and Immunology, 5640 Fisher's Lane, Rockville, Maryland 20852, USA

HEPATITIS delta virus (HDV) is a subviral human pathogen that requires hepatitis B virus (HBV) for packaging^{1,2}. Concurrent infection by HBV and HDV increases the risk of severe liver disease compared to infection with HBV alone³. The HDV genome is a closed circular RNA of about 1,700 bases which is replicated through an RNA intermediate, the antigenome⁴. Both RNAs can be folded into highly base-paired, rod-shaped structures, similar to the plant viroid RNAs⁵. Two forms of the sole HDV protein, hepatitis delta antigen, are derived from a single open reading frame by RNA editing; the enzymes responsible for the editing have not been characterized. Here we report that the purified enzyme dsRAD (for double-stranded-RNA-adenosine deaminase) can edit HDV antigenomic RNA *in vitro*. Most important, we observe that mutations in critical sequences of the antigenome have identical effects on *in vitro* and *in vivo* editing, suggesting that dsRAD, or a closely related enzyme, is responsible for editing HDV RNA *in vivo*.

The larger form of hepatitis delta antigen, HDAg-p27, is formed when RNA editing converts an amber stop codon (UAG) to a tryptophan codon (UGG), extending the open reading frame by 19 amino acids⁶. The smaller form, HDAg-p24, is required for replication⁷, whereas HDAg-p27 represses replication⁸ and is required for packaging^{1,2}; thus, RNA editing plays a central role in the HDV life cycle. Typically, 20–50% of isolated virions are edited at the site referred to here as the amber/W site (ref. 9 and J.L.C., manuscript in preparation).

Editing has been proposed to occur on the genomic RNA, where a U-to-C change could produce the requisite codon change in the messenger RNA (Fig. 1)^{10,11}. However, more recent work shows that editing occurs on the antigenome¹², which requires editing of an adenosine. Although editing of the antigenomic A to a G could explain the codon change, the double-stranded character of HDV suggested that dsRAD might be responsible for RNA editing of HDV (a model is shown in Fig. 1). dsRAD converts adenosine within dsRNA to inosine (I), which, like G, prefers to pair with C^{13,14}. In support of the idea that dsRAD edits

HDV, a base-paired structure is required for editing¹¹, and the U that is 5' of the amber/W site is consistent with dsRAD's 5' nearest neighbour preference (A = U > C > G)¹⁵.

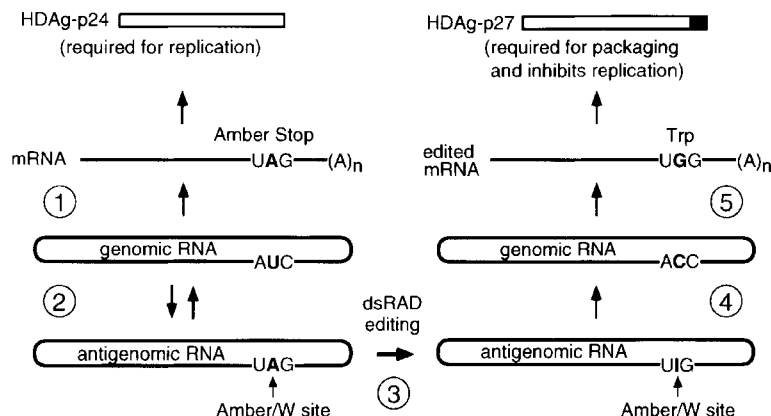
We directly tested the hypothesis that dsRAD can edit HDV by incubating *in vitro* synthesized antigenomic RNA with dsRAD purified from *Xenopus laevis* eggs. Editing at the amber/W site was monitored by digesting polymerase chain reaction (PCR)-amplified complementary DNA with restriction endonucleases. PCR products from unedited RNA cannot be cut by *StyI* or *DsaI*, whereas editing leads to PCR products that can be digested. When HDV antigenomic RNA was incubated with increasing amounts of purified dsRAD, editing increased proportionally, as indicated by the appearance of the two shorter restriction fragments (Fig. 2a). This result was surprising because to date it has been impossible to develop *in vitro* RNA-editing systems that consist solely of purified components¹⁶. As shown, *in vivo* levels of editing (20–50%) were easily observed (Fig. 2a, b).

The antigenomic RNA was radiolabelled with ³²P-ATP during transcription so that we could also monitor total adenosine-to-inosine conversion by thin-layer chromatography. As shown in Fig. 2a, b, inosine production also increased proportionally with the amount of dsRAD added. Consistent with the known properties of dsRAD^{17,18}, double-stranded RNA, but not single-stranded RNA, was a competitor in the deamination reaction and editing at the amber/W site (Fig. 2b).

The amount of A-to-I conversion observed in the experiments shown in Fig. 2 indicated that sites other than the amber/W site were deaminated. For example, of the 297 adenosines contained in the antigenomic RNA used for the experiment in Fig. 2b, 0.6%, or an average of 1.8 adenosines per antigenome, were deaminated. Interestingly, cDNAs made from antigenomic RNA derived from HDV passaged in woodchucks, in addition to an A-to-G change at the amber/W site, show other A-to-G changes which could correspond to additional A-to-I changes in the RNA¹⁹. These cDNAs also show U-to-C conversions, suggesting that dsRAD deaminates genomic RNA, and that the changes are passed to the antigenome during replication.

To determine more precisely the number of deamination events occurring in our *in vitro* reaction, and whether genomic RNA could also be deaminated, we reacted full-length genomic and antigenomic RNA in quintuplicate and assayed the amount of editing and A-to-I conversion (data not shown). Genomic RNA, which was not edited at the amber/W site, did show $0.5 \pm 0.1\%$ A-to-I conversion. The antigenomic RNA sample showed $16 \pm 1\%$ editing and $0.81 \pm 0.05\%$ A-to-I conversion. Each RNA had ~340 adenosines, so the observed amount of inosine predicts an average of 2.7 modifications per antigenome and 1.7 modifications per genome. Our prediction of about 3 deaminations per antigenome was substantiated by sequencing 45 cloned cDNAs made from a small region (358 nucleotides) of the *in vitro* modified antigenomic RNA. This analysis also showed that 17% of the clones contained

FIG. 1 A model for RNA editing of HDV. (1) Replication-competent genomes are transcribed to produce an mRNA encoding HDAg-p24. (2) HDAg-24 enables replication of the genome by RNA polymerase II (ref. 4), generating antigenomic RNA. (3) dsRAD acts on antigenomic RNA to convert the adenosine at the amber/W site to an inosine. (4) Like G, inosine prefers to pair with C; thus, after replication, the genome has a C at the amber/W site instead of a U. (5) The edited genome is transcribed to yield an mRNA encoding HDAg-p27. HDAg-p27, which contains a 19-amino-acid extension (shaded), inhibits replication and helps packaging of the HDV genome by HBV surface antigen.



A-to-G changes at the amber/W site, consistent with the value of 16% predicted by our restriction enzyme assays. The location of A-to-G changes other than the amber/W site showed some overlap with A-to-G changes found *in vivo*, but also many differences. At present it is unclear if these differences are meaningful because only a single RNA sample was sequenced both in the *in vivo* studies¹⁹ and the *in vitro* studies reported here. However, if dsRAD is responsible for multiple deamination sites within HDV, one might expect differences between *in vitro* and *in vivo* sites because changes that interfere with the HDV life cycle may be selected against during successive passages.

The minimal number of deaminations occurring *in vitro* was notable, given that the amount of dsRAD in our *in vitro* reaction was sufficient to deaminate >50% of the adenosines in an 800-base-pair completely duplexed RNA (CAT dsRNA; not shown). However, the latter substrate, as well as most of the artificial RNAs previously used to study dsRAD, did not contain the structural features (bulges, mismatches and loops) found in HDV; as proposed, dsRAD may use such structural features to increase selectivity in its biological substrates^{15,20}.

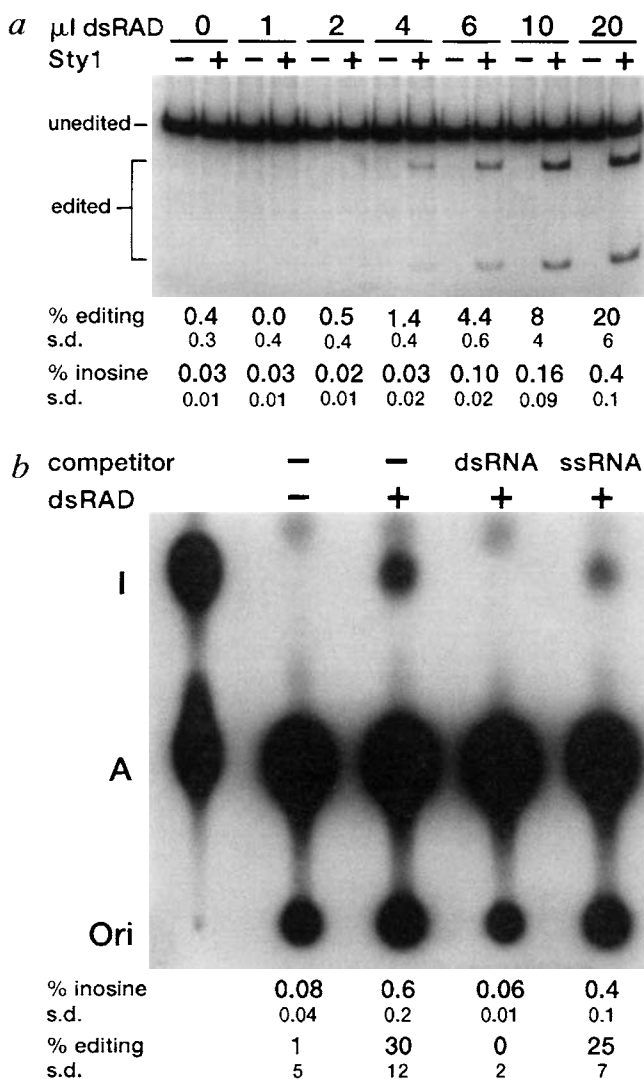
Although the features of *in vitro* editing of HDV shared similarities with *in vivo* editing, we wished to obtain further evidence that editing in our *in vitro* reaction was related to editing of HDV *in vivo*. Previous studies have shown that *in vivo* editing can be altered by mutations near the amber/W site¹¹. Thus, we directly compared the effect of previously characterized muta-

tions on editing *in vitro* and during replication in transfected cells (Fig. 3). In all five of the mutants tested, the effect of the mutation on *in vitro* editing was identical to its effect on editing in cells. As might be expected, the effect of each sequence change on editing was consistent with known properties of dsRAD¹⁵. For example, mutations that altered the thermodynamic stability of base-paired regions surrounding the editing site (M9, M10, M9/10) also altered editing, and a mutation that created a new editing site with a G as its 5' neighbour (M4) was a poor target for dsRAD¹⁵. Mutation of the A·C mismatch at the amber/W site to an A·U pair (M6) also reduced editing. This behaviour is similar to that found at the mismatched R/G editing site of certain glutamate-receptor pre-mRNAs²¹. Although not yet proven, glutamate receptor mRNA is the only other RNA for which existing data support *in vivo* editing by dsRAD²²⁻²⁴. The observation that nucleotide changes near the editing site had identical effects on *in vivo* editing and editing *in vitro* with pure dsRAD supports the hypothesis that the editing we observe *in vitro* is related to the editing of HDV RNA *in vivo*.

We have established an *in vitro* system for HDV RNA editing that can mimic the efficiency of *in vivo* editing and also exhibits similar sequence requirements near the editing site. Our data strongly implicate dsRAD, or a closely related protein, as the enzyme that edits HDV RNA *in vivo*. Although definitive proof that dsRAD edits HDV RNA *in vivo* will require further studies, perhaps involving inactivation of dsRAD by gene targeting, the *in*

FIG. 2 Purified dsRAD can edit HDV. *a*, Increasing amounts of dsRAD lead to increasing amounts of editing, as evidenced by StyI digestion of PCR products to yield fragments of 374 and 131 bp. Antigenomic RNA (10 fmol, 1/3Δ 1mer) was reacted for 1 h with dsRAD in a total volume of 80 μl, and editing and inosine production quantified. Quantification was based on relative radioactivity. Numbers below the autoradiogram show the average per cent editing (% PCR product digested), and inosine ((I/(A + I + origin)) × 100; *n* = 3). *b*, TLC of mononucleotides derived from antigenomic RNA. Antigenomic RNA (10 fmol, 1/3Δ 1.2mer) was reacted for 2 h with or without 15 μl dsRAD and, as indicated, 20 fmol competitors (CAT RNA, as in ref. 20). Numbers for triplicate reactions (40 μl each) are as in *a*. Unlabelled lane shows 5'AMP and 5'IMP markers.

METHODS. HDV RNAs were a unit length of genome or antigenome (1mer), or slightly longer (1.2mer); except when designated as full-length, RNAs contained deletions that removed 1/3 of the RNA but maintained the rod-shaped structure (1/3Δ). The deletions had no observable effect on editing or inosine production. The following plasmids, described previously¹², were linearized with *Bam*HI, and transcribed to make the RNAs listed: 1/3Δ 1.2mer antigenomic RNA (pCMV3-DC-Δ1 × 1.2), 1/3Δ 1mer antigenomic RNA (pCMV3-DCA1 × 1), 1/3Δ 1.2mer RNAs M4 (1011U/1012C), M6 (580A), M9 (1014C), M10 (578G) and M9/10 (578G/1014C) (derivatives of pCMV3-DC-Δ1 × 1.2). pGDC-1¹¹ was linearized with *Ecl*136II or *Spe*I and transcribed to produce full-length 1mer antigenomic or genomic RNA, respectively. RNAs were purified on denaturing gels as described¹⁵, except gels were 6% polyacrylamide. The dsRAD preparation showed a single band of 120K by SDS-PAGE and was identical to that used in a recent study²⁰; specific activity, 0.6 pmol inosine per μl dsRAD per h with CAT dsRNA. Reactions with dsRAD were in standard assay buffer²⁵ except RNasin was 160 U ml⁻¹. After incubation at 25 °C, reactions were processed as described²⁵. Editing assays were as described¹² using primer pairs 5414 and 6657, or for full-length HDV, 5414 and 5415; a single-cycle PCR radiolabelling reaction was performed after 35 PCR cycles. Radiolabelled product (5 μl) was digested with StyI or *Dsa*I, and analysed as described¹². Sequencing of cDNAs showed no deaminations that would interfere with restriction enzyme digestion. TLC assays on RNA (8 fmol) were as described²⁶. Full-length 1mer antigenomic or genomic RNAs (10 fmol) were incubated with dsRAD (15 μl) at 25 °C for 2 h (total volume, 40 μl).



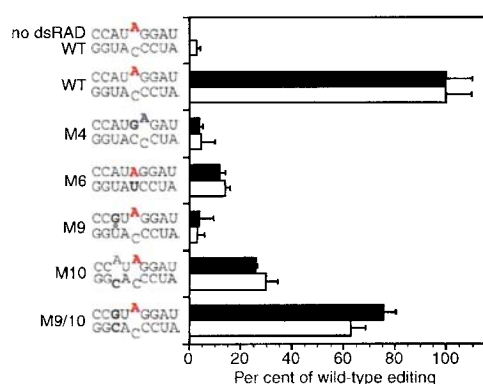


FIG. 3 *In vitro* editing has similar sequence requirements as editing *in vivo*. The graph compares *in vitro* (open bars) and *in vivo* (filled bars) editing for 5 HDV mutants. Bar length reflects per cent editing normalized to wild-type (100%). Antigenomic sequences surrounding the editing site are shown (1,016–1,008 paired with 584–576) with mutated nucleotides (bold), the amber/W site (red) and a potential new editing site created in mutant M4 (blue) indicated. Mutations were restricted to those that did not alter the encoded HDAG amino-acid sequence, and to maintain replication, to those that, in the absence of editing, maintained the stop codon at the amber/W site. Virion production, which is critically dependent on editing, was sharply reduced for all mutants except M9/10 (J.L.C., data not shown). Before normalization, wild-type editing *in vivo* was $25 \pm 3\%$ or $21 \pm 4\%$ as assayed with *Styl* or *Dsal*, respectively, and *in vitro*, $44 \pm 4\%$ (*Styl*) and $46 \pm 3\%$ (*Dsal*).

METHODS. For *in vitro* reactions (20 μ l; $n = 4$), 5 fmol antigenomic RNA (1/3 Δ 1.2mer) was incubated with dsRAD (7.5 μ l) for 2 h. For editing in cells, cDNA constructs that expressed replication-competent HDV RNA were transfected into HuH7 cells, and RNA collected 12 days post-transfection as described¹¹ ($n = 3$ except for M6 where $n = 2$). RNAs edited *in vitro* or *in vivo* were assayed using *Styl* (wt, M4, M6, M10) or, because some mutations destroyed the *Styl* site, *Dsal* (wt, M9, M9/10).

in vitro system described here may represent the first example of an RNA editing reaction to occur *in vitro* with purified components. **Note added in proof:** Another *in vitro* reaction involving purified dsRAD has now been described²⁷. □

Received 4 December 1995; accepted 2 February 1996.

- Wu, J. C. et al. *J. Virol.* **65**, 1099–1104 (1991).
- Ryu, W. S., Bayer, M. & Taylor, J. J. *J. Virol.* **66**, 2310–2315 (1992).
- Rizzetto, M. *Hepatology* **3**, 729–737 (1983).
- Lai, M. M. C. A. *Rev. Biochem.* **64**, 259–286 (1995).
- Wang, K. S. et al. *Nature* **323**, 508–514 (1986).
- Luo, G. X. et al. *J. Virol.* **64**, 1021–1027 (1990).
- Kuo, M. Y., Chao, M. & Taylor, J. J. *J. Virol.* **63**, 1945–1950 (1989).
- Chao, M., Hsieh, S. Y. & Taylor, J. J. *J. Virol.* **64**, 5066–5069 (1990).
- Tang, J. R. et al. *J. med. Virol.* **42**, 1–6 (1994).
- Zheng, H., Fu, T. B., Lazinski, D. & Taylor, J. J. *J. Virol.* **66**, 4693–4697 (1992).
- Casey, J. L., Bergmann, K. F., Brown, T. L. & Gerin, J. L. *Proc. natn. Acad. Sci. U.S.A.* **89**, 7149–7153 (1992).
- Casey, J. L. & Gerin, J. L. *J. Virol.* **69**, 7593–7600 (1995).
- Bass, B. L. & Weintraub, H. *Cell* **55**, 1089–1098 (1988).
- Wagner, R. W., Smith, J. E., Cooperman, B. S. & Nishikura, K. *Proc. natn. Acad. Sci. U.S.A.* **86**, 2647–2651 (1989).
- Polson, A. G. & Bass, B. L. *EMBO J.* **13**, 5701–5711 (1994).
- Bass, B. L. in *The RNA World* (eds Gesteland, R. & Atkins, J.) 383–418 (Cold Spring Harbor Laboratory Press, New York, 1993).
- Bass, B. L. & Weintraub, H. *Cell* **48**, 607–613 (1987).
- Wagner, R. W. & Nishikura, K. *Molec. cell. Biol.* **8**, 770–777 (1988).
- Netter, H. J. et al. *J. Virol.* **69**, 1687–1692 (1995).
- Hurst, S. R., Hough, R. F., Aruscavage, P. J. & Bass, B. L. *RNA* **1**, 1051–1060 (1995).
- Lomeli, H. et al. *Science* **266**, 1709–1713 (1994).
- Rueter, S. M., Burns, C. M., Coode, S. A., Mookherjee, P. & Emerson, R. B. *Science* **267**, 1491–1494 (1995).
- Yang, J.-H., Sklar, P., Axel, R. & Maniatis, T. *Nature* **374**, 77–81 (1995).
- Melcher, T., Maas, S., Higuchi, M., Keller, W. & Seeburg, P. H. *J. biol. Chem.* **270**, 8566–8570 (1995).
- Hough, R. F. & Bass, B. L. *J. biol. Chem.* **269**, 9933–9939 (1994).
- Saccomanno, L. & Bass, B. L. *Molec. cell. Biol.* **14**, 5425–5432 (1994).
- Melcher, T. et al. *Nature* **379**, 460–464 (1996).

ACKNOWLEDGEMENTS. We thank R. Hough and S. Hurst for providing purified dsRAD and T. Brown for technical assistance. This work was supported by funds to B.L.B. from the NIH, the Pew Scholars Program, and the David and Lucile Packard Foundation, and by funds to J.L.C. from the NIH. B.L.B. is an HHMI assistant investigator.

CORRESPONDENCE and requests for materials should be addressed to B.L.B. (e-mail: bas@tulip.med.utah.edu).

A new pattern for helix–turn–helix recognition revealed by the PU.1 ETS-domain–DNA complex

Ramadurgam Kodandapani*, Frédéric Pio*,
Chao-Zhou Ni*, Gennaro Piccialli†,
Michael Klemsz‡, Scott McKercher*,
Richard A. Maki*§ & Kathryn R. Ely*

* La Jolla Cancer Research Center at The Burnham Institute,
10901 North Torrey Pines Road, La Jolla, California 92037, USA

† Dipartimento di Chimica Organica e Biologica, Università degli Studi di
Napoli Federico II, 80134 Napoli, Italy

§ Neurocrine Biosciences, San Diego, California 92121, USA

THE Ets family of transcription factors, of which there are now about 35 members^{1,2}, regulate gene expression during growth and development. They share a conserved domain of around 85 amino acids³ which binds as a monomer to the DNA sequence 5'-C/AGGAA/T-3'. We have determined the crystal structure of an ETS domain complexed with DNA, at 2.3-Å resolution. The domain is similar to $\alpha + \beta$ (winged) 'helix–turn–helix' proteins and interacts with a ten-base-pair region of duplex DNA which takes up a uniform curve of 8°. The domain contacts the DNA by a novel loop–helix–loop architecture. Four of the amino acids that directly interact with the DNA are highly conserved: two arginines from the recognition helix lying in the major groove, one lysine from the 'wing' that binds upstream of the core GGAA sequence, and another lysine, from the 'turn' of the 'helix–turn–helix' motif, which binds downstream and on the opposite strand.

The PU.1 [*Sp1-1*, *Spf1-1*] transcription factor is an Ets protein expressed in haematopoietic cells^{4,5}. PU.1 is a regulatory protein for differentiation of monocytes and macrophages and for B-cell maturation (reviewed in ref. 2). The ETS domain of PU.1 was co-crystallized with a 16 base-pair oligonucleotide containing the recognition sequence⁶. The structure was solved by the multiple isomorphous replacement and anomalous scattering (MIRAS) method (Table 1). The electron density was clearly defined (Fig. 1) for residues 171 to 258, which encompasses the entire conserved ETS domain. The PU.1 domain assumes a tight globular structure ($33 \times 34 \times 38 \text{ Å}^3$) formed by three α -helices and a four-stranded antiparallel β -sheet (Fig. 1). The domain topology is similar to the structures of other Ets family proteins Fli-1 (ref. 7), murine Ets-1 (ref. 8) and human Ets-1 (ref. 9) determined in solution by NMR. The structural studies revealed a common folding pattern for ETS domains that is similar to $\alpha + \beta$ helix–turn–helix (HTH) DNA-binding proteins including CAP¹⁰ and resembles 'winged' HTH proteins such as GH5 (ref. 11), HNF-3 γ (ref. 12) and HSF (ref. 13). There are three sites of protein–DNA contact: the recognition helix ($\alpha 3$), the loop between β -strands 3 and 4 (a 'wing') and the turn in the HTH motif ($\alpha 2$ –turn– $\alpha 3$). The turn between $\alpha 2$ and $\alpha 3$ is longer than the equivalent in many other HTH proteins, and is actually a loop. The DNA-binding motif in PU.1, and probably other members of the Ets family, can be described more appropriately as a loop–helix–loop motif. Therefore the large Ets family defines a new variant subclass of the helix–turn–helix DNA-binding proteins with a novel mode of DNA recognition.

The protein–DNA contacts in the PU.1 complex are detailed in Fig. 2. Four strictly conserved residues on the surface of the domain are likely to be important for DNA binding by all members of the Ets family. Arg 232 and Arg 235 emanate from helix $\alpha 3$ and contact bases in the GGAA sequence in the major groove. These contacts represent the core structure for DNA

‡ Present address: Department of Microbiology and Immunology, Indiana University School of Medicine, Indianapolis, Indiana 46202-5120, USA.

recognition by members of the Ets family because they involve both strictly conserved amino acids and bases in the consensus sequence recognized by these transcription factors (see Fig. 3b). The equivalent arginines 81 and 84 in Ets-1 (ref. 9) do not contact the GGAA bases, but intermolecular nuclear Overhauser effects between these arginines and DNA were observed in the Fli-1 NMR studies⁷. Lys 245 extends from $\beta 3$ just adjacent to the loop ('wing'), and Lys 219 is located in the 'loop' of the HTH motif. Lys 245 contacts the phosphate backbone of the GGAA strand in the minor groove upstream from the core sequence (Fig. 3c) and Lys 219 forms a salt bridge with the phosphate backbone of the opposite strand downstream of the GGAA core (Fig. 3d). Substitutions of glycine at each of these four conserved sites abolished DNA binding, confirming the functional importance of these contacts (see Fig. 2).

Mutations of conserved residues that contact the phosphate backbone also affect DNA binding. Substitution of glycine at Leu 174 or Trp 215 abolished DNA binding in PU.1. Similarly, substitution of any amino acid in Ets-1 (ref. 14) at the equivalent of PU.1 residues Lys 219 and Arg 222 that bind the phosphate

backbone disrupted DNA binding. These minor-groove contacts might represent a conserved pattern for protein 'docking' in the Ets family. In Fli-1 (ref. 7), the equivalents of Leu 174, Lys 219 and Lys 222 showed large chemical shifts on DNA binding in the NMR studies (the counterpart of Trp 215 was buried).

Water molecules also participate in protein-DNA recognition in the PU.1 complex (Fig. 2). There are 27 well-ordered solvent molecules around the DNA. Solvent molecules in the major groove are hydrogen-bonded to the bases and also form a hydrogen-bonded network between the two strands that might contribute to the stability of the duplex and consequently influence specific DNA recognition. Conserved Arg 232 and Arg 235 each form direct and water-mediated contacts with the bases. Three other residues also contact DNA bases through water molecules: Thr 226, Gln 228 and Asn 236. These residues are not conserved in the Ets family and might represent interactions that are unique to the PU.1 protein. Thr 226 and Gln 228, at the amino-terminal end of helix $\alpha 3$, make water-mediated contacts with bases C25 and C26 respectively that are base-paired to guanines 8 and 9 in the core sequence.

TABLE 1 Structure determination and refinement

	Native	Hg	I (29)	I (13)	I (31)
Phasing statistics					
Resolution (Å)	2.3	3.0	2.9	3.0	2.8
Observed reflections	60,095	25,081	20,709	20,512	23,308
Unique reflections	20,105	14,902	13,258	12,910	15,397
Completeness (%)	97	79	65	69	68
R_{sym} (%) [*]	5.0	3.6	4.0	4.3	3.6
R_{iso} (%) to 3.0 Å†		13.0	14.4	15.9	13.0
Number of sites		2	2	2	2
For isomorphous data ($I/\sigma \geq 3$)					
Phasing power‡		1.33	1.76	1.04	0.98
To resolution (Å)		3.0	3.0	3.0	3.0
R_{Cullis} §		0.62	0.57	0.68	0.67
For anomalous data ($I/\sigma \geq 3$)					
Phasing power		1.0	1.41	1.13	1.43
To resolution (Å)		3.0	3.0	3.0	3.0
Mean figure of merit (10–3.0 Å) is 0.65.					
Refinement statistics					
Resolution range		8–2.3 Å			
Average B (Å ²)		20.1			
Crystallographic R -factor (%)		23.7			
R_{free} (%) ¹⁶		29.9			
Number of reflections used		16,898 $F > 3\sigma(F)$			
Number of protein atoms		1,486			
Number of DNA atoms		1,300			
Number of solvent atoms		88			

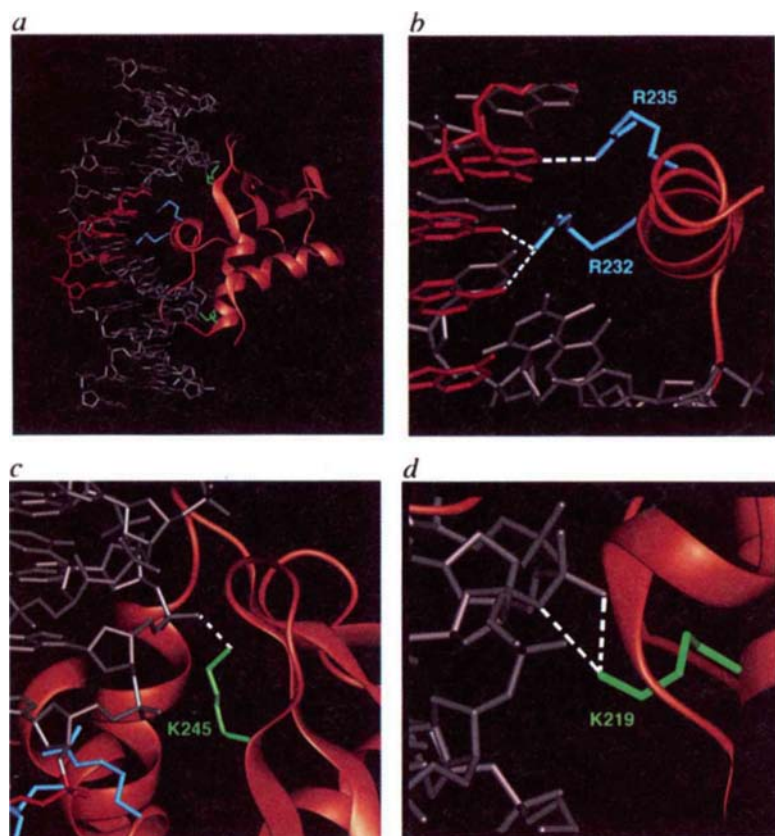
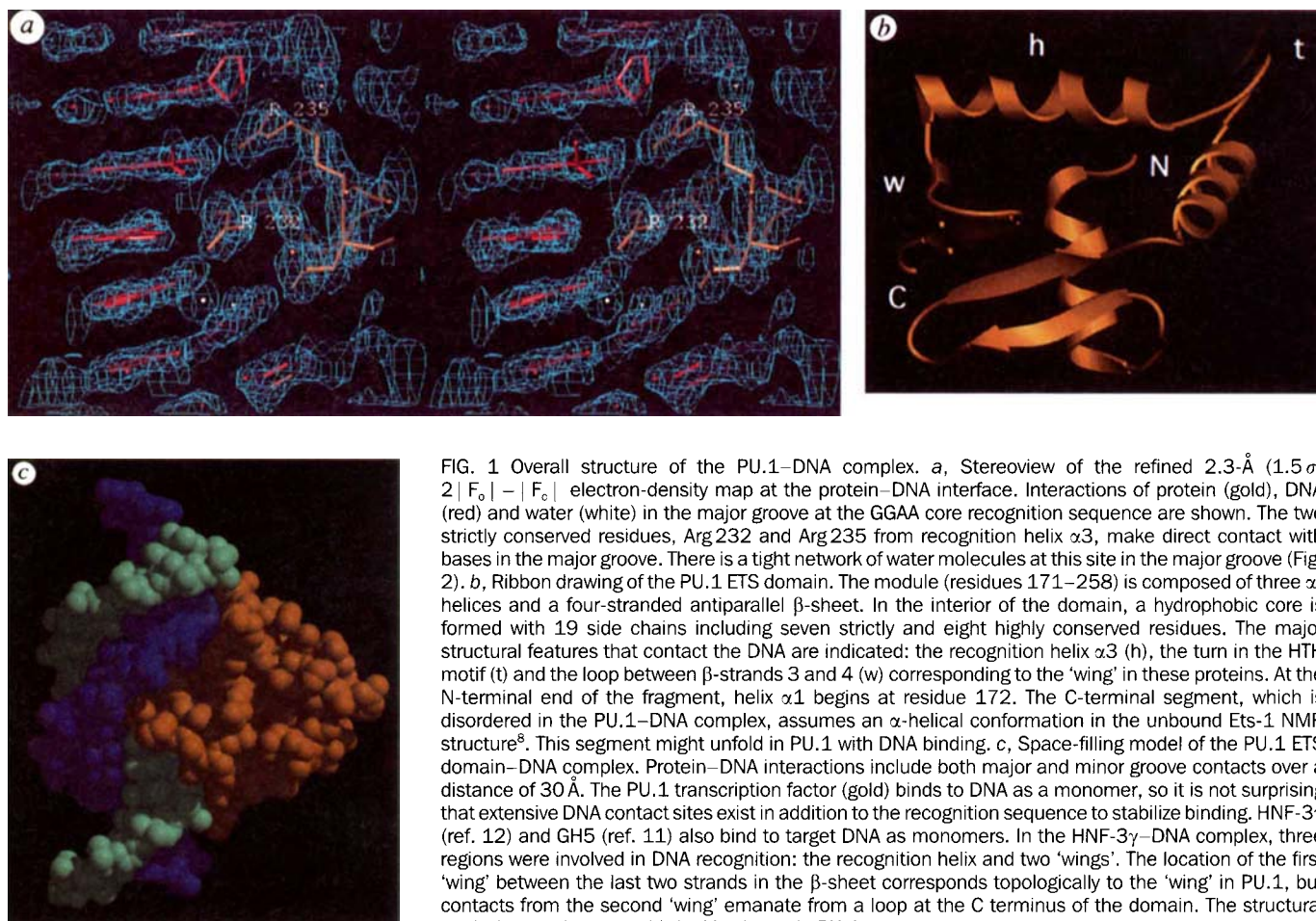
The crystallization of the PU.1 ETS domain (residues 160–272) with a 16-bp synthetic DNA oligonucleotide containing the recognition sequence was described previously⁶. Crystals formed in the space group C_2 with $a = 89.1$, $b = 101.9$, $c = 55.6$ Å and $\beta = 111.2^\circ$, with two complexes in the asymmetric unit. **Phase determination.** Four heavy-atom derivatives were prepared by soaking crystals of the native complex and by co-crystallizing iodinated oligonucleotides with the PU.1 domain. The locations of the iodinated bases are indicated in Fig. 2. Multiple isomorphous replacement phases, including anomalous data, were calculated. The package PHASES¹⁷ was used to refine heavy-atom positions, B -factor/occupancies and to calculate phases to 3.0-Å resolution with an overall figure of merit of 0.65. The initial MIRAS map (3.0 Å) was improved by solvent flattening by the method of Wang¹⁸ and with non-crystallographic density averaging. **Model building and refinement.** The improved MIRAS electron-density map was used to build the model with the interactive graphics programs TOM based on FRODO¹⁹ and O²⁰. The density for the DNA helix was a prominent feature of the map. To fit the DNA, an 'ideal' B-DNA duplex was generated with the program QUANTA (Molecular Simulations, Inc.) and fitted to the density as a rigid body. After the DNA was positioned, a polyaniline chain was constructed with the BONES option of the Alberta/Caltech program TOM. Subsequently side chains for all residues with clear electron density were added to the model. There were 11 disordered residues at the N terminus of the domain and 14 disordered residues at the C terminus so these amino acids were not included in the model. For all other residues representing the complete ETS domain, the electron density was clear (see Fig. 1) and allowed unambiguous fitting of both backbone and side-chain atoms. Manual adjustments of individual DNA bases were made to fit the electron density. In the program X-PLOR²¹, the stereochemistry of the protein was optimized to bond and angle parameters developed by Engh and Huber²² and for DNA by using parameters of Parkinson *et al.*²³. Weak restraints were placed on all ribose conformations. One cycle of simulated annealing at 3,000 K (ref. 24) was followed by cycles of manual model building, positional refinement and B -factor refinement. More data were added as the refinement progressed in increments: 3, 2.8, 2.6 and 2.3 Å. A total of 88 solvent oxygens ($\langle B \rangle = 22$ Å²) have been added to the model at this stage of the refinement. Main-chain torsion angles for all non-glycine residues fall within energetically favourable Ramachandran boundaries²⁵. The r.m.s. difference for 84 α -carbon atoms in the two complexes in the asymmetric unit is 0.35 Å.

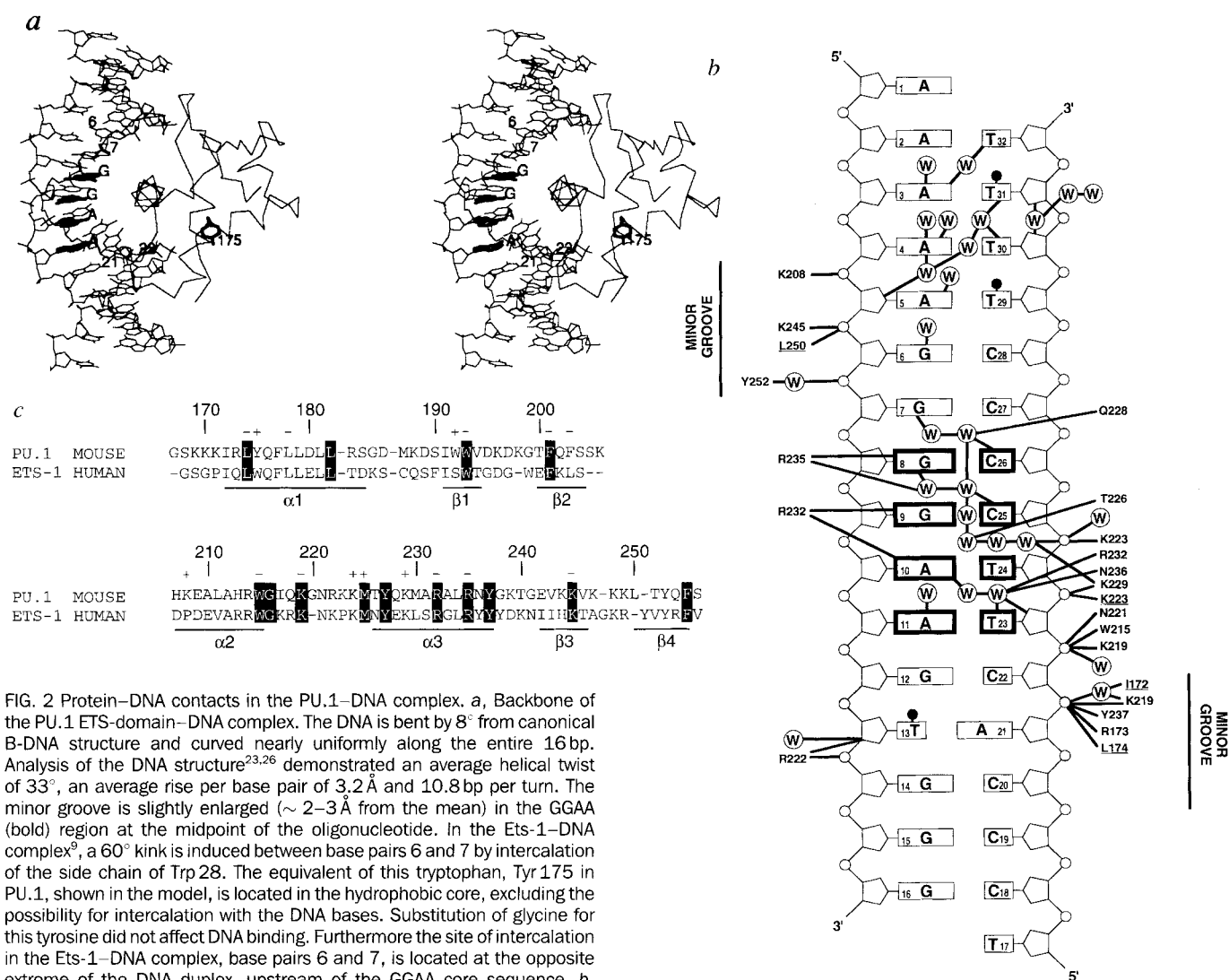
^{*} R_{sym} is $\sum |I - \langle I \rangle| / \sum \langle I \rangle$.

† R_{iso} is $\sum ||F_{\text{PH}}| - |F_P|| / \sum |F_P|$, where $|F_P|$ and $|F_{\text{PH}}|$ are structure factors for the protein and derivative, respectively.

‡ Phasing power is the r.m.s. value of $|F_H|/E$, where E is residual lack of closure.

§ R_{Cullis} is $\sum ||F_{\text{PH}}| \pm |F_P| - |F_{\text{H(calc)}}|| / \sum |F_{\text{PH}} - F_P|$ for centric reflections, where $F_{\text{H(calc)}}$ is the calculated heavy-atom structure factor.





The turn in the HTH motif is actually a loop, and because the sequences in this loop as well as the loop ('wing') between strands $\beta 3$ and $\beta 4$ are not strictly conserved among members of the Ets family, these residues might be important sites for specific recognition by individual members of the family. In fact, the lengths of both of the contact loops differ among members of the family, with the PU.1 loop containing an 'extra' glycine at residue 220 and lacking a glycine after residue 247. Such conformational differences are expected between family members, but the contrast between the PU.1 and Ets-1 complexes was unexpected. The striking distinction in the mode of DNA contact by the PU.1 and Ets-1 domain could reflect extreme evolutionary divergence between members of the Ets family. Alternatively, it should be noted that the Ets-1-DNA complex was formed under denaturing conditions^{9,15} and it is possible that the Trp intercalation occurred early during the renaturation step with subsequent protein refolding.

the PU.1 and Ets-1 ETS domains, representing extremes of evolutionary divergence in the family. Residues strictly conserved in all Ets proteins are shown in black boxes; dashes indicate gaps within the family. Numbering and secondary structural features of the PU.1 domain are indicated. The results of mutational analysis when glycine was substituted for a residue are also shown. The effects of the interchanges are labelled + or - above the sequence, indicating that DNA binding was retained or abolished. Mutations were generated essentially as described²⁷.

Future extensive mutational studies of amino acids that contact DNA in Ets proteins are needed to identify residues that mediate recognition of a specific DNA sequence by a given family member. Ultimately, crystal structures of other Ets proteins complexed to DNA must be compared to distinguish unique DNA contacts. □

Received 12 January; accepted 19 February 1996.

1. Wasyluk, B., Hahn, S. L. & Giovane, A. *Eur. J. Biochem.* **211**, 7–18 (1993).
2. Moreau-Gachelin, F. *Biochim. biophys. Acta* **1198**, 149–163 (1994).
3. Karim, F. et al. *Genes Dev.* **4**, 1451–1453 (1990).
4. Kierns, M. J., McKeircher, S. R., Celada, A., Van Beveren, C. & Maki, R. A. *Cell* **61**, 113–124 (1990).
5. Moreau-Gachelin, F., Mattei, M. G., Tambourin, R. & Tavittian, A. *Oncogene* **4**, 1449–1456 (1989).
6. Pio, F. et al. *J. Biol. Chem.* **270**, 24258–24263 (1995).
7. Liang, H. et al. *Nature Struct. Biol.* **1**, 871–875 (1994).
8. Donaldson, L. W., Petersen, J. M., Graves, B. J. & McIntosh, L. P. *EMBO J.* **15**, 125–134 (1996).
9. Werner, M. H. et al. *Cell* **83**, 761–771 (1995).
10. Schultz, S. C., Shields, G. C. & Steitz, T. A. *Science* **253**, 1001–1007 (1991).

11. Ramakrishnan, V., Finch, J. T., Graziano, V., Lee, P. L. & Sweet, R. M. *Nature* **362**, 219–223 (1993).
12. Clark, K. L., Halay, E. D., Lai, E. & Burley, S. K. *Nature* **364**, 412–420 (1993).
13. Harrison, C. J., Bohm, A. A. & Nelson, H. C. M. *Science* **263**, 224–227 (1994).
14. Mavrothalassitis, G., Fisher, R. J., Smyth, F., Watson, D. K. & Papas, T. S. *Oncogene* **9**, 425–435 (1994).
15. Werner, M. H., Clore, G. M., Gronenborn, A. M., Kondoh, A. & Fisher, R. J. *FEBS Lett.* **345**, 125–130 (1994).
16. Brünger, A. T. *Nature* **355**, 472–475 (1992).
17. Furey, W. & Swaminathan, S. *Am. crystallogr. Ass. Meeting* **18**, 73 (1990).
18. Wang, B. C. *Methods Enzymol.* **115**, 90–112 (1985).
19. Jones, T. A. *Methods Enzymol.* **115**, 157–171 (1985).
20. Jones, T. A., Zhou, J., Cowan, S. W. & Kjeldgaard, M. *Acta crystallogr.* **A47**, 110–119 (1992).
21. Brünger, A. T. *X-PLOR Manual Version 3.1* (Yale Univ. Press, New Haven, CT, 1992).

22. Engh, R. A. & Huber, R. *Acta crystallogr.* **A57**, 392–400 (1991).
23. Parkinson, G., Vojtechovsky, J., Clowney, L., Brünger, A. T. & Berman, H. M. *Acta crystallogr.* **D52**, 57–64 (1996).
24. Brünger, A. T., Krokowski, A. & Erickson, J. W. *Acta crystallogr.* **A46**, 585–593 (1990).
25. Ramachandran, G. N. & Sasiekharan, V. *Adv. Protein Chem.* **23**, 283–438 (1968).
26. Babcock, M. S. & Olson, W. K. *J. molec. Biol.* **237**, 98–124 (1994).
27. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. & Pease, L. R. *Gene* **77**, 51–59 (1989).

ACKNOWLEDGMENTS. This work was supported by grants from the U.S. Army and NIH. R.K. and F.P. made equal contributions to this study. We thank J. Knight and R. Mitchell for synthesis and purification of DNA oligonucleotides, M. Hasham for excellent graphics illustrations, and K. Riddle-Hilde for preparing the manuscript for publication.

CORRESPONDENCE AND MATERIALS. Requests to be addressed to K.R.E.

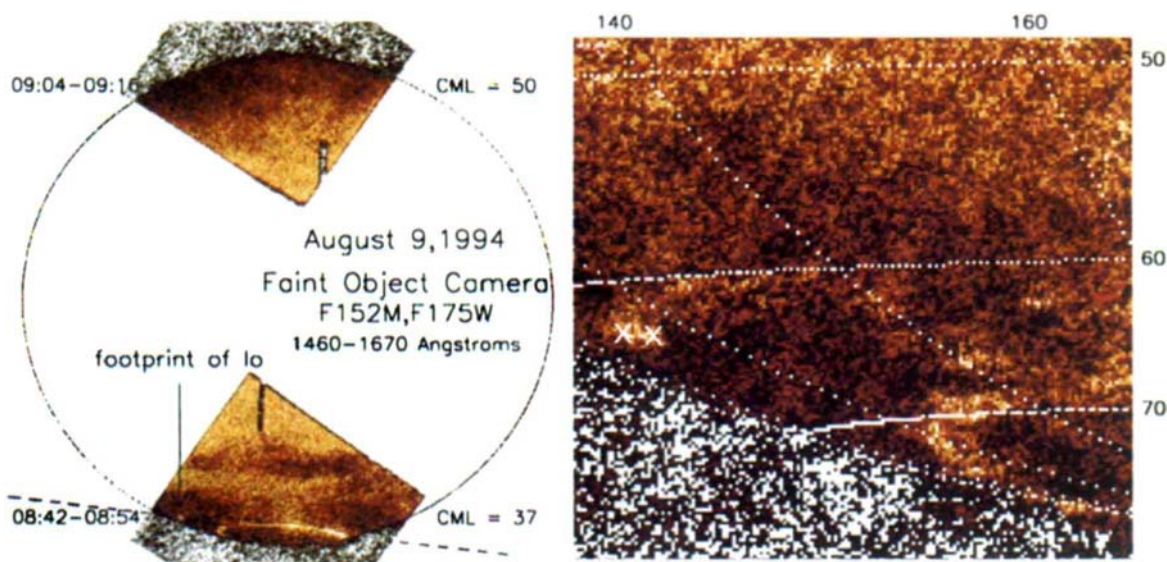
CORRECTION

Rapid energy dissipation and variability of the Io–Jupiter electrodynamic circuit

Renée Prangé, Daniel Rego, David Southwood, Philippe Zarka, Steven Miller & Wing Ip

Nature **379**, 323–325 (1995)

IN Fig. 1 of this Letter, a dashed line was omitted that was referred to in the legend to Fig. 3a. The corrected figure is shown here.



Images of the FUV Jovian aurorae and of the far-ultraviolet signature of the Io magnetic flux tube with the Hubble Space Telescope Faint Object Camera (f/96). The plot in Fig. 3a was taken along the dashed line in the left image, across the south Io footprint and the auroral oval.

KNOW YOUR COPY RIGHTS RESPECT OURS

The publication you are reading is protected by copyright law. Photocopying copyright material without permission is no different from stealing a magazine from a newsagent, only it doesn't seem like theft.

If you take photocopies from books, magazines and periodicals at work your employer should be licensed with CLA. Make sure you are protected by a photocopying licence.



The Copyright Licensing Agency Limited
90 Tottenham Court Road, London W1P 0LP
Telephone: 0171 436 5931 Fax: 0171 436 3986

High-resolution carbohydrate profiling by capillary gel electrophoresis

Andras Guttman

Capillary gel electrophoresis with laser-induced fluorescent detection has been found to be a powerful separation and characterization tool for fluorescently labelled carbohydrates.

INCREASING evidence has shown that carbohydrates on glycoproteins are often important as recognition determinants in receptor–ligand or cell–cell interactions, in the modulation of protein folding and in the regulation of protein bioactivity¹. Additionally, complex carbohydrates are important in providing energy and structural supports for cells. Changes in the biological activity of glycoproteins are often associated with alterations in protein glycosylation either through site variations or changes in the structure of the oligosaccharide occupying a particular site. Until recently, various chromatographic techniques, and high-concentration polyacrylamide gel electrophoresis were mostly used to separate carbohydrates^{2,3}. In the latter instance, the sugar molecules were labelled either with an ultraviolet active dye or a fluorescent dye, with charged groups for their electric-field-mediated mobilization³.

Separation of complex carbohydrates with high resolution by capillary electrophoresis is a new, rapidly developing area⁴. This novel, high-performance separation method in conjunction with laser-induced fluorescence (LIF) detection is an optimal separation platform for profiling fluorophore-labelled oligosaccharides. It provides fast and efficient separations, a high degree of automation and multiple injections from microlitre-scale sample volumes. One benefit of capillary electrophoresis with LIF detection is its sensitivity (low attomol range) and high separation efficiency (greater than one million theoretical plates per metre).

The analytical system described here for carbohydrate profiling is based on the use of capillary gel electrophoresis to separate and analyse oligosaccharides released from glycoproteins. It is necessary to use neutrally coated quartz capillary columns in order to suppress electroosmotic flow that would significantly decrease the reproducibility and increase the time of

analysis at the optimal pH of 4.75. All the capillary gel electrophoresis experiments were performed with the eCAP asparagine-linked oligosaccharide profiling kit on a P/ACE 5500 capillary electrophoresis system (both from Beckman Instruments, Fullerton, California). The separations were monitored on column with a Beckman LIF detection system using a 4-mW argon-ion laser. The tem-

perature of the capillary in the instrument was controlled at $20\text{ }^{\circ}\text{C} \pm 0.1\text{ }^{\circ}\text{C}$. The glycoproteins of bovine fetuin and bovine pancreatic ribonuclease B were purchased from Sigma (St Louis, Missouri). The oligosaccharide profiling of glycoproteins begins with the release of the glycans by enzymatic or chemical means. N-linked (asparagine-bound) glycans are usually released by using the enzyme peptide *N*-glycosidase F (PNGase F, Glyko, Novato, California), O-linked glycans are cleaved off the protein by hydrazinolysis. The released oligosaccharides are then labelled with a novel fluorophore of 8-aminopyrene-1,3,6-trisulphonate (APTS) at their reducing terminus by reductive amination⁵, so that only one molecule of APTS is attached to each oligosaccharide molecule. The fluorophore labelling efficiency is greater than 97 per cent, without any bias in labelling selectivity⁶. The optimal excitation wavelength of the APTS-labelled oligosaccharides is close to 488 nm, the wavelength of the Ar-ion laser. The wavelength of the maximum emission of the fluorophore–sugar conjugates is 520 nm. The APTS derivatives of oligosaccharides are readily separated (profiled) by capillary gel electrophoresis.

The separation of an APTS-labelled maltooligosaccharide ladder standard is shown in Fig. 1. The numbers above the peaks of the upper trace represent the degree of polymerization. The enzymatically (PNGase F)-released asparagine-linked oligosaccharides from bovine fetuin are represented by the middle trace. The high resolution of the maltoligosaccharides ($R_s > 5$) suggested that the method could resolve various closely related oligosaccharides having very fine structural differences, for example positional isomers. The two doublets in the middle trace of Fig. 1 (fetuin) were postulated and identified by coinjection, as tetrasialo-triantennary (F1 and F2) and trisialo-triantennary (F3 and F4) structures, respectively⁷. The individual peaks within each doublet have a different number of $\alpha 2 \rightarrow 3$ and $\alpha 2 \rightarrow 6$ linked sialic acids. The structures of the four peaks are shown in Fig. 2. It is seen in Fig. 1 that the electrophoretic mobilities of the two doublets of F1/F2 and F3/F4 (having 15 and 14 sugar units, respectively) are similar to those of the maltotetraose and maltopentaose in the ladder (with four and five sugar units, respectively). This phenomenon could be explained by the presence of sialic acid

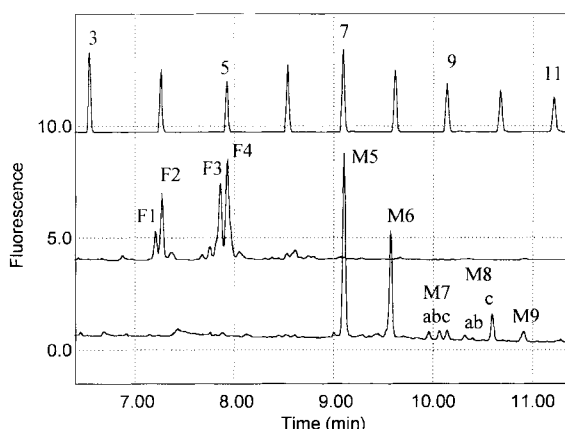


FIG. 1 Electropherograms of APTS-labelled glycans of bovine fetuin (middle trace) and bovine ribonuclease B (lower trace) compared with the malto-oligosaccharide ladder standard (upper trace). Numbers on the upper trace correspond to the degree of polymerization of the glucose oligomers. F1, tetrasialo-triantennary- $2\alpha 2,6$; F2, tetrasialo-triantennary- $2\alpha 2,3$; F3, trisialo-triantennary- $2\alpha 2,6$; F4, trisialo-triantennary- $2\alpha 2,3$; M5–M9, mannose 5 – mannose 9. Conditions, capillary, 40-cm (effective) neutrally coated capillary (eCAP NCHO) with 50- μm inner diameter; buffer, 25-mM acetate buffer, pH 4.75; detection, laser-induced fluorescence; excitation, 488 nm; emission, 520 nm; applied field strength, 500 V cm^{-1} ; temperature, $20\text{ }^{\circ}\text{C}$.

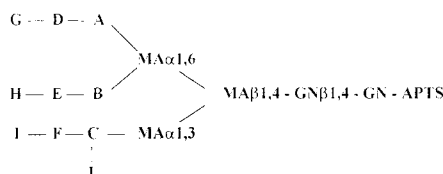
perature of the capillary in the instrument was controlled at $20\text{ }^{\circ}\text{C} \pm 0.1\text{ }^{\circ}\text{C}$. The glycoproteins of bovine fetuin and bovine pancreatic ribonuclease B were purchased from Sigma (St Louis, Missouri).

The oligosaccharide profiling of glycoproteins begins with the release of the glycans by enzymatic or chemical means. N-linked (asparagine-bound) glycans are usually released by using the enzyme peptide *N*-glycosidase F (PNGase F, Glyko, Novato, California), O-linked glycans are

Peaks	A	B	C	D	E	F	G	H	I	J
F1	GN β 1,2	GN β 1,2	GN β 1,4	GA β 1,4	GA β 1,4	GA β 1,4	SA α 2,6	SA α 2,6	SA α 2,3	SA α 2,6
F2	GN β 1,2	GN β 1,2	GN β 1,4	GA β 1,4	GA β 1,4	GA β 1,4	SA α 2,3	SA α 2,6	SA α 2,3	SA α 2,6
F3	GN β 1,2	GN β 1,2	GN β 1,4	GA β 1,4	GA β 1,4	GA β 1,4	SA α 2,6	SA α 2,6	SA α 2,3	-
F4	GN β 1,2	GN β 1,2	GN β 1,4	GA β 1,4	GA β 1,4	GA β 1,4	SA α 2,3	SA α 2,6	SA α 2,3	-
M5	MA α 1,6	MA α 1,3	-	-	-	-	-	-	-	-
M6	MA α 1,6	MA α 1,3	MA α 1,2	-	-	-	-	-	-	-
M7a	MA α 1,6	MA α 1,3	MA α 1,2	-	MA α 1,2	-	-	-	-	-
M7b	MA α 1,6	MA α 1,3	MA α 1,2	-	-	MA α 1,2	-	-	-	-
M7c	MA α 1,6	MA α 1,3	MA α 1,2	MA α 1,2	-	-	-	-	-	-
M8a	MA α 1,6	MA α 1,3	MA α 1,2	MA α 1,2	MA α 1,2	-	-	-	-	-
M8b	MA α 1,6	MA α 1,3	MA α 1,2	-	MA α 1,2	MA α 1,2	-	-	-	-
M8c	MA α 1,6	MA α 1,3	MA α 1,2	MA α 1,2	-	MA α 1,2	-	-	-	-
M9	MA α 1,6	MA α 1,3	MA α 1,2	MA α 1,2	MA α 1,2	MA α 1,2	-	-	-	-

FIG. 2 Schematic representation of the oligosaccharide structures (right) separated by capillary gel electrophoresis from the bovine fetuin (F1, F2, F3 and F4) and the ribonuclease B (M5-M9) glycan pool after APTS labelling. Symbols, GN, N-acetyl-glucosamine; SA, sialic acid; GA, galactose; MA, mannose. α , β and numbers near symbols refer to the position of linkages.

Guttman, Table 1



residues that results in high charge-to-mass ratio for these structures⁷. The APTS-labelled oligosaccharides have three sulphonic acid groups, thus three negative charges. As the sialic acid residues are almost completely dissociated at pH 4.75 ($pK_{\text{sialic acid}} \approx 2.6$), they impart either four (F1 and F2) or three (F3 and F4) additional negative charges on these molecules, resulting in the observed much higher migration velocity. It is important to note that at pH 2.5 the doublets could not be separated⁷.

The released and labelled high-mannose type carbohydrates of bovine ribonuclease B (Fig. 1, lower trace) were identified by coinjection with authentic oligosaccharides⁸. Here, the branched high-mannose structures of M5-M9 (Fig. 2), having seven to eleven sugar units, respectively, migrate with velocities similar to that of maltoheptaose to malto-undekaose. The slight discrepancy in their electrophoretic mobilities was probably caused by the coiling tendency of the linear malto-oligomers, which results in higher hydrodynamic volumes as compared with the branched, and therefore less flexible, high-mannose structures⁸. This may explain the excellent separations of the three positional isomers of the mannose-7 (M7a;b;c) and mannose-8 (M8a;b;c) oligosaccharides attained (Fig. 2).

Capillary gel electrophoresis with LIF detection has been found to be a sensitive and powerful technique for separation and characterization of fluorescently labelled carbohydrates. The detection limits provided by injecting only a few nanolitres of samples are in the low attomole range. Consequently, oligosaccharides released

from only a few micrograms of glycoproteins can be quickly and accurately profiled. The reproducibility of the absolute migration times with this methodology was

found to be $RSD < 1.2$ per cent, and the relative to maltose, $RSD < 0.09$ per cent, making possible easy peak assignments in the released glycan pools for high-precision run-to-run and day-to-day operations. This allows easy applicability and validation of this new method for the modern biotechnology industry for rapid, sensitive and high-resolution analysis of oligosaccharides, released from recombinant glycoproteins. □

Andras Guttman is with Beckman Instruments, 2500 Harbour Blvd, Fullerton, California, 92634-3100. The author acknowledges the stimulating discussions with Csaba Horváth of Yale University, and the support of this work by Nelson Cooke and Sandy Brunet of Beckman Instruments. For more information fill in reader enquiry number 100.

- Varki, A. *Glycobiology* **3**, 97-130 (1993).
- Townsend, R. R. in *Carbohydrate Analysis* (ed. El Rassi, Z.) (Elsevier, Amsterdam, 1995).
- Jackson, P. *Analyt. Biochem.* **216**, 243-252 (1994).
- Chiesa, C. & Horváth, Cs. *J. Chromatogr.* **645**, 337-352 (1993).
- Evangelista, R., Liu, M. S. & Chen, F. T. A. *Analyt. Chem.* **67**, 2239-2245 (1995).
- Guttman, A., Chen, F. T. A., Evangelista, R. & Cooke, N. *Analyt. Biochem.* **233**, 234-242 (1996).
- Guttman, A., Evangelista, R. & Chen, F. T. A. *Electrophoresis* **17**, 412-417 (1996).
- Guttman, A. & Pritchett, T. *Electrophoresis* **16**, 1906-1911 (1995).

Cellular connection

Some recent product developments for the cell biologist include a liposomal transfection reagent, an optical biosensor system and more.

PERSEPTIVE BIOSYSTEMS has introduced a new methodology for the **automated mapping of peptides**, which is said to reduce analysis time, improve reproducibility and provide higher protein recovery for the characterization of trace protein amounts. The technique uses Porosyme immobilized enzyme cartridges on the company's integral micro-analytical workstation for the rapid chemical modification of proteins, on-line protein digestion and reversed-phase chromatographic peptide mapping. PerSeptive recently published a paper on automated isolation, peptide mapping and mass spectrometric sequencing of an entire haemoglobin protein from serum. According to the manufacturer, this approach was able to discriminate one of 574 amino acids that was different between normal and sickle cell haemoglobin in two hours instead of several days.

Reader Enquiry No. 101

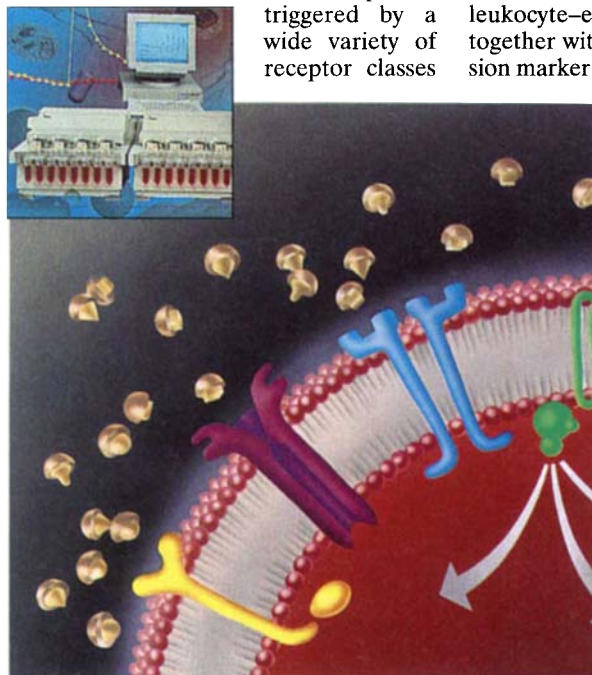
AutoFill Expert, the automatic, **programmable filling module** is designed for reliable and economical filling of different-

sized sample vessels. Sample vessels can be arranged in lines or in a honeycombed pattern. The unit is programmable for up to 99 racks on a filling area of 420 mm × 245 mm. The dosing pump is fully programmable with an LCD display and variable settings for dispensing volumes from one millilitre to ten litres. It is compatible with the Compupump and Dosifit peristaltic pumps, which can be operated with AutoFill through an interface cable. Compupump is designed for higher flexibility and Dosifit focuses on efficiency and economy.

Reader Enquiry No. 102

Molecular Devices offers an 8-page, 4-colour brochure that describes the Cytosensor **microphysiometer**. This technical brochure contains 14 different experimental data graphs demonstrating the use of the system for a variety of applications: agonist and antagonist profiling, G-protein differentiation, signalling specificity, apoptosis, cytotoxicity, signal transduction elucidation, T-cell signalling, *in vitro* toxicity, and microbial sus-

ceptibility and challenge assays. In addition, three examples are given of how the Cytosensor System is able to detect novel receptor responses not seen with traditional assay methods. The microphysiometer is a single-assay system that can detect responses triggered by a wide variety of receptor classes



Molecular Devices' brochure covers agonist/antagonist profiling, G-protein differentiation, signalling specificity, apoptosis, cytotoxicity, signal transduction and more.

coupled to different second messengers. These receptor classes include G-protein-linked receptors, ligand-gated ion channels, and tyrosine-kinase receptors activated via growth factors, cytokines, and neurotrophic factors.

Reader Enquiry No. 103

Apoptosis

Now available from BioWhittaker for the study of apoptosis is the Annexin V detection kit, which detects onset of apoptotic cell death process. This should prove to be important for those in cancer, AIDS and autoimmune research. The company claims that the kit is the only direct marker to detect cell death before the loss of membrane integrity. Traditional techniques detect apoptosis towards the terminal stage of the process, characterized by breakdown of the cell membrane and total disruption of the cell. However, according to the manufacturer Annexin V allows researchers to probe the origins of the apoptotic process offering quick and reliable detection and quantification of the onset of cell death. The kit can detect phosphatidylserine, a phospholipid normally only found on the inside of intact cell membranes.

Reader Enquiry No. 104

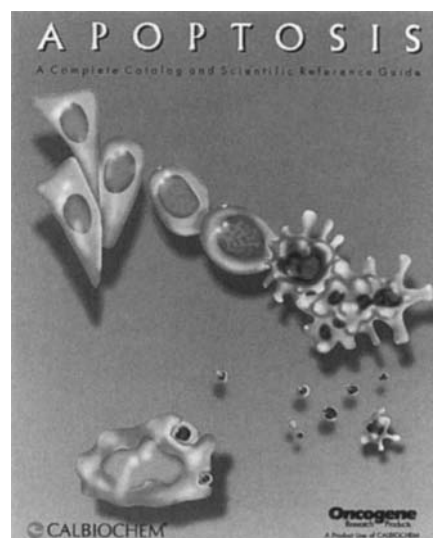
Serotec offers a range of **monoclonal antibodies against cell-adhesion markers**. To complement this they have now released a new wall chart, which complements their CD/antigen expression wall chart. The new full-colour wall chart details all of the major leukocyte-leukocyte and leukocyte-endothelial cell interactions, together with details of all of these adhesion marker antibodies, many of which are active functionally in blocking and stimulation assays. In addition to their panel of reagents against human integrins, immunoglobulin superfamily members and selectins, the company also has a wide range of antibodies recognizing these markers in mouse and rat. Many of these are available conjugated to FITC and to red phycoerythrin — for performing functional assays, many azide-free formats are also available.

Reader Enquiry No. 105

Bender offers a number of **new reagents for studying apoptosis**. Among these are APO-1/Fas (CD95), which can induce apoptosis upon tagging with the monoclonal antibody SM1/1 (BMS138); the antibody clone SM1/23 (BMS140), which is able to inhibit SM1/1-induced apoptosis; APO-1/Fas, expressed on a broad variety of normal and malignant cells and on activated human B and T cells; and SM1/17 (BMS139), which does not induce apoptosis, but is recommended for staining procedures of living cells. APO also occurs in a soluble form (sAPO-1) devoid of a transmembrane region. Elevated sAPO-1 levels are present in sera from patients with high- and low-grade malignant B- and T-cell leukaemias and systemic lupus erythematosus, which can be detected by ELISA (BMS230). Annexin V (BMS306) is also available which specifically binds phosphatidylserine. This is said to make it a fast and reliable tool for the detection and quantification of apoptosis on a single-cell basis.

Reader Enquiry No. 106

Oncogene Research Product's 200-page **catalogue and scientific handbook, Apoptosis**, contains over 300 antibodies, assays, peptides, proteins and probes for the study of apoptosis. Products include regulators of apoptosis, p53 and related products, cell-cycle and proliferation markers, and specific inducers and inhibitors of apoptosis. Eileen White has written a mini-review outlining the



Oncogene Research products catalogue.

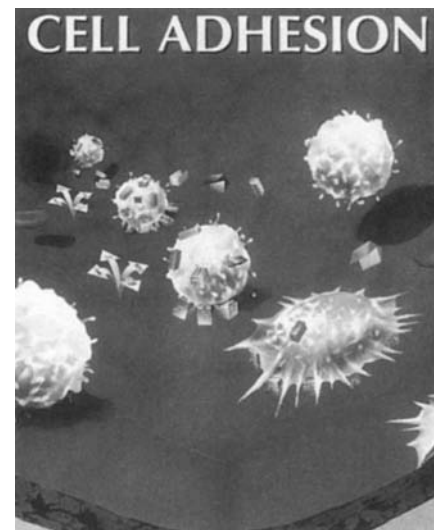
major proteins and pathways involved in programmed cell death. All Oncogene products are well referenced, and a comprehensive methods and protocol section is included.

Reader Enquiry No. 107

Cell adhesion

Human recombinant proteins for coating microtitre plates in cell-adhesion studies are now available from R&D Systems in quantities ranging from two hundred micrograms to hundreds of milligrams. The company offers sE-selectin, sL-selectin, sP-selectin and sVCAM-1. Coming soon are sICAM-1 and sCD31 (PECAM). According to the manufacturer, these proteins are >97 per cent pure, are free endotoxin and are biologically active. They can be used in adhesion screening assays to investigate blockers of adhesion, such as peptide antagonists, mimetic drugs or antibodies.

Reader Enquiry No. 108



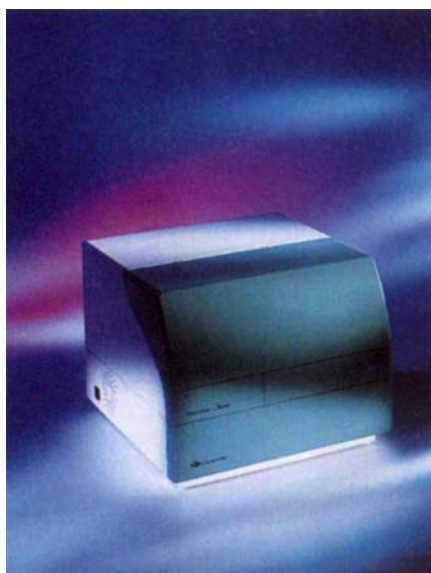
R&D Systems' coated microtitre plates.

Promega offers non-radioactive proliferation and cytotoxicity assays that **measure cell proliferation, viability and death**. These have been developed to eliminate the need for hazardous radioisotopes and to reduce significantly the number of steps involved in the procedure. Promega's CellTiter 96 and CellTiter 96 AQueous assay kits rely on the conversion of a tetrazolium salt to an intensely coloured, formazan-dye product within living cells. Cell death is also easily measured using the Cytotox 96 kit, which detects the release of the cytosolic enzyme lactate dehydrogenase from lysed cells.

Reader Enquiry No. 109

Bioanalytical systems

Labsystems has launched Fluoroskan Ascent, a system intended for **advanced fluorescence assays**. This optical system is made up of a highly focused light beam and a new accurate fluorometric reader. It also offers high scanning resolution with over 3,500 measurement positions per plate, as well as an excitation wavelength that starts at 320 nm, allowing for measurement in the ultraviolet range. This new fluorometric reader can use an extensive variety of plate configurations and it can be switched easily from above-



Fluoroskan Ascent from Labsystems.

to below-plate reading. To speed up reaction times and improve performance, a built-in orbital shaking feature with continuously adjustable speed and diameter, has also been added.

Reader Enquiry No. 110

A new optical biosensor system, designed for higher throughput, unattended **bio-molecular interaction analysis**, has been announced by Affinity Sensors. IAsys Auto+ offers all the advantages of the company's stirred micro-cuvette system and choice of fully or semi-automated operation, as well as simultaneous analysis of two channels for higher throughput. 'Intelligent' software is aimed at higher productivity and mirrors the user logic of interactive manual operation by giving the option to analyse binding responses automatically and to adapt sample contact times or experimental sequences. Using this 'adaptive sequence control' (ASK) capability, the new biosensor allows method programs to include data interrogation, simple mathematical formulae, and to write if/then commands. Alternatively, immediate hands-on operation of the instrument, with direct control of the autosampler, is provided by unique image-based software known as 'point and transfer' (PAT). The instrument uses a biosensor with two vibro-stirred, micro-cuvette wells to allow simultaneous experiments using different ligands, analyses and chemistries, with no risk of cross-contamination. Direct pipetting into the micro-cuvette wells is designed to minimize sample usage and avoid the risk of blockage presented by complex fluids. Micro-cuvettes with a choice of derivatized sensor surfaces combined with well-established coupling procedures provide a variety of ligand immobilization strategies for a wide range of molecule types.

Reader Enquiry No. 111

The Merck-Hitachi **reaction system 655A-13** was conceived to enable up to three different reaction solutions to be pumped into the eluate for detection of the reaction products and subsequent post-column derivatization. All parts of the pump that come into contact with solution are made of inert materials so that aggressive reaction solutions can be used. Up to three reaction capillaries can be connected. All conventional post-column derivatization procedures, such as the analysis of amino acids, biogenic amines and polyamines, sugars, carbamate pesticides and glyphosate can be carried out with the system and automated with ease. A number of thermostatable reaction capillaries are available as optional accessories.

Reader Enquiry No. 112

Cell-handling systems

Eppendorf has announced the introduction PatchMan 5173, a new **micromanipulator for use in performing patch-clamp electrophysiology recordings**. A pushbutton keyboard, joystick control and stepper motors combine to provide the speed, range, precision, stability and convenience required in



Eppendorf's PatchMan micromanipulator.

whole-cell, single-channel and nystatin electrophysiology experiments. PatchMan is an addition to the existing ECET line of micromanipulators and microinjection instruments, which are currently used to inject proteins or DNA into individual cells.

Reader Enquiry No. 113

A new scientific workstation from Melles Griot has been engineered for maximum **vibration isolation** to ensure high instrument resolution and repeatability in sensitive applications. The company has translated its experience as a manufacturer of research-grade optical tables to produce a range of vibration isolation workstations that are suitable for supporting patch-clamp apparatus, cell-injection and micromanipulation equipment, interferometers, microscopes, non-contact surface measurement devices, and other motion-sensitive instruments. The basic workstation consists of a welded steel frame, with a self-levelling, pneumatic damping system, which provides both vertical and horizontal isolation for the 60-mm-thick instrument platform. A stainless steel honey-

ADVERTISEMENT

New Prices

Oligonucleotides
no setup charge

from **\$1.17** per base,

VISA or Mastercard payment

0.2 microM scale

including gel picture and synthesis report

Small to Large Scale Synthesis

All modifications available



Specializing in Safe Handling of Large Volume
Orders

with prompt and courteous service.

For further information call:

1-800-249-7841

Phone: 1-610-993-93 23

Fax: 1-610-993-94 64

E-Mail: ORDERS@DNAAGENCY.COM

Distributor in Japan: Biotech Japan, Fax: 81

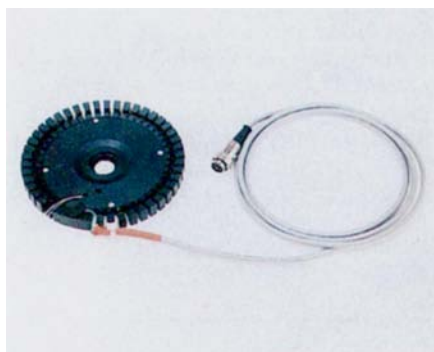
886 87 3061

READER ENQUIRY NO. 661

comb-core construction is designed to ensure that the platform has maximum rigidity and minimum vibration transmittance. A range of accessories include rear- and side-shelf systems to support ancillary equipment and a front support rail with adjustable padded arm rests. All accessories rest on independent supports attached to the base frame and are therefore isolated from the instrument platform. These sophisticated workstations are available in a variety of sizes and with optional 110-mm-thick, steel or aluminium instrument platforms with plain, roll-formed or clean-room grade surfaces. A range of six standard sizes is available.

Reader Enquiry No. 114

Medical Systems has introduced a new **micro-incubation system used to maintain patch or tissue slices on the stage of a microscope**. The Patch Slice micro-incubation system comprises a versatile micro-incubation chamber (PSMI), an innovative coverslip dish (PS-CSD), and a bipolar temperature controller (TC-202). This system is said to provide a complete solution to the challenge of electrophysiological study of synaptic slice connections at physiological temperatures using low-noise, whole-cell patch recording either with or without water immersion of the microscope objective in the recording medium. This micro-incubator is a shallow, annular assembly, which surrounds the central slice holding dish. Two independent channels of room-temperature perfusant are delivered to the incubator, which in turn delivers them to the dish at a precisely controlled temperature. A



Medical Systems' micro-incubation set up.

separate gas inlet provides temperature-controlled gas across the dish. The TC-202, connected electrically to the micro-incubator, provides noise-free control of the desired system temperature (10–15 °C below ambient to 25 °C above ambient) using the micro-incubator warming/cooling plate as the control. The PS-CSD consists of two 22-mm diameter coverslips held in a circular frame. The slice sits on the top coverslip with the immersion objective reaching the fluid surface. The bottom coverslip prevents condensation that would otherwise dete-

riorate illumination by the microscope condenser below. To enable placement of the patch electrode, entry angles as small as 15°, with respect to the horizontal, can be achieved for slice access.

Reader Enquiry No. 115

Assays

Cell Census Plus is a new **cell proliferation assay** designed to allow researchers to measure simultaneously cell proliferation and to study phenotypically distinct subsets of a mixed-cell population using flow cytometry. Available from Sigma Chemical the new system may be used on all cell types without affecting the cell's biological or proliferative function and requires no radioactive materials. The system uses an aliphatic reporter molecule (PKH26), which acts as a fluorescent membrane dye, and a patented diluent that allows the dye to incorporate strongly and evenly into the cell's membranes. The PKH26 dye's rapid uptake and stable retention in the plasma membrane and equal distribution into membranes of the daughter cells make it suitable for cell tracking and measuring proliferation. The system can also be used with immunogating techniques to measure proliferation of a select subset of cells within the whole proliferating cell culture. Analysis of data is performed using Modfit and special software, which interprets raw histogram data and provides quantitative data on the number of cell divisions (up to 10 generations), relative abundance of each daughter generation, the fraction of non-proliferating cells, and the proliferation index. To optimize assay parameters, PKH26 reference microbeads are included to standardize set up of the flow cytometer, ensuring parallel parameters for all assays and normalized cell counts.

Reader Enquiry No. 116

For expedient and efficient **measurement of matrix metalloproteases (MMP)**, Amersham has a new range of Biotrak MMP ELISAs. These focus on cellular communication by measuring MMPs, specifically, total MMP-1 (also known as interstitial collagenase), total TIMP-1 (tissue inhibitor of metalloproteinases-1) and the activated MMP-1/TIMP-1 complex. Their defined specificity is said to enable more information to be obtained; the microtitre plate allows for ease of use and enhanced precision. Using highly specific antibodies precoated onto microtitre plates, the new ELISAs are said to allow measurement in 4.5 or 5.5 h, as compared with several days for a traditional bioassay. Conventional bioassays, such as zymography, measure an enzyme's degradative ability; for example, in the case MMP-1, on collagen. However, this requires the enzyme to be activated and,

in addition, the natural inhibitor of MMP-1, TIMP-1, is often present, which means that the measured activity is not a true reflection of the enzyme concentration. This type of assay is also relatively nonspecific — it can distinguish neither MMP-1 and MMP-8 nor the inhibitors TIMP-1, -2, -3 and α_2 -macroglobulin. According to the manufacturer, the MMP ELISAs overcome all of these problems and offer a faster assay protocol.

Reader Enquiry No. 117

Molecular means

A new, low-cost, **DNA purification system for cells and tissues** is now available from the Nest Group. The NucleoSpin C & T system can be used for the extraction of 20–40 µg genomic DNA from body fluids and tissue samples, as well as for the determination of bacteria. Each micro-spinning cup can be closed and fits standard micro-centrifuge tubes. The procedure can be completed in 20 min or less. The plasmid DNA is eluted in pH 9.0 TE and is ready for a variety of manipulations, including restriction analysis, sequencing, amplification or bacterial transformation.

Reader Enquiry No. 118

For high-efficiency **transfection of eukaryotic cells**, Boehringer Mannheim offers its new Dosper liposomal transfection reagent. Designed to transfect a wide range of cell lines in the presence or absence of serum (for example, CHO-K1 cells), it eliminates the need for serum deprivation. This polycationic preparation binds negatively charged DNA more effectively than monocationic reagents. Bottled under argon to prevent oxidation

ADVERTISEMENT

PACEMAKER®

Custom peptide, conjugate and antibody services

When your research depends on high quality at low cost turn to **PACEMAKER®**

All **PACEMAKER®** peptides, whether >70%, >80% or >95% purity, undergo hplc and mass spectrometric analysis.

Peptides are available in quantities from 1mg to 100mg and, with prices from £115 per peptide and £90 per conjugate, even the tightest budget is catered for.

Most importantly, you may come to us for advice. Based on many years experience and expertise we will try to find the best approach to fulfilling your custom peptide and antibody requirements.

Understanding the needs of the Life Scientist

A F F I N I T I

Research Products Limited

Mamhead Castle, Mamhead
Exeter EX6 8HD, United Kingdom
Tel +44(0) 1626 891010 Fax +44(0) 1626 891090

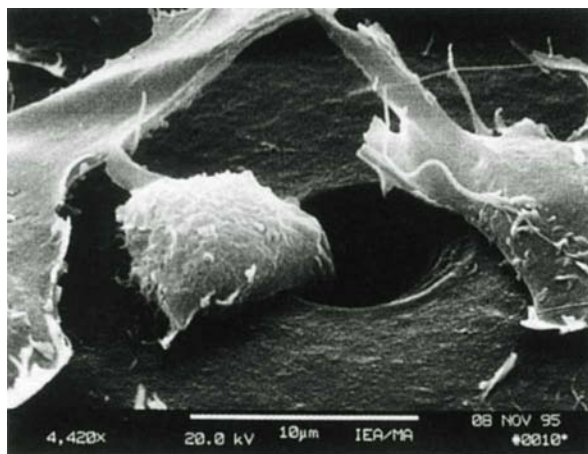
READER ENQUIRY NO 70

and peroxide formation, Dosper possesses ester linkages that are easily metabolized by cellular esterases, avoiding the cytotoxicity common to ether-linked reagents.

Reader Enquiry No. 119

Miscellaneous

Becton Dickinson Labware has recently introduced a broad range of BioCoat cell environment products and SMC STIM specialty media for the culture of vascular smooth muscle cells (SMC). The culture system that supports SMC differentiation *in vitro* incorporates a serum-free or low-serum D-STIM differentiation medium and GFR Matrigel matrix plates with reduced levels of growth factors and an optimized concentration of matrix. The



Becton Dickinson Labware SMC STIM specialty media for the culture of vascular smooth muscle cells.

culture system that induces rapid growth in SMC contains a low-serum P-STIM proliferation medium and BioCoat collagen I plates. Finally, SMC migration or invasion can be studied with the use of the GFR matrigel invasion chamber that contains 8-µm membrane pores occluded with a thin layer of a reconstituted basement membrane with reduced levels of

growth factors (GFR Matrigel matrix). The cells are allowed to migrate for 24 h in the presence of serum-free SMC M-STIM migration medium and a chemotactic agent (PDGF-BB). These cell environment products and SMC STIM specialty media can be used to construct physiologically relevant *in vitro* systems to study SMC in their differentiated (contractile), proliferative (synthetic), or invasive migratory phenotypic states.

Reader Enquiry No. 120

The *Oxidative Metabolism Handbook*, a new 20-page booklet from Electrochemistry Separations Analysis (ESA), provides a background to the science of **reactive oxygen species (ROS)**; its role in neurological disorders, aging, cancer, apoptosis and more; the mechanisms of oxidative damage, defence mechanisms, assays, drugs and diseases associated with free radicals. Divided into sections, this monograph gives the reader a greater understanding of oxidative stress (when the production of damaging ROS overwhelms the body's defensive mechanisms, its prevention and consequences when the body's natural defenses fail). Specific HPLC-EC based methods and brief overviews of analytical methodologies and articles relevant to the ROS field are included.

Reader Enquiry No. 121

Life Technologies, producer of Gibco BRL products, has introduced the Green Lantern-1 Vector, a new use for green fluorescent protein (GFP) as a **transfection reporter vector for expression in mammalian cells**. The result is a fluorescent signal and faster visualization. This product has several mutations that improve fluorescence over 40fold compared with wild-type GFP.

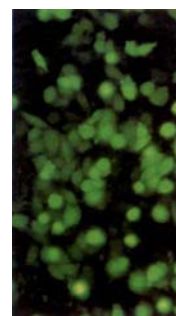
To enhance detection, a Ser 65→Thr mutation was employed to increase signal intensity up to sixfold, in addition to shifting the excitation peak to 490 nm. Compared with wild-type GFP, this shift permits more efficient detection of the GFP-expressing cells in FACS analysis, as well as fluorescence microscopy. Over 80 mutations introduced into the gene create a pattern of codon usage that reflects the abundance of tRNAs in a mammalian population. This "humanized" gene is translated more efficiently and increases the level of GFP protein in the cell by up to 22fold. The Green Lantern-1 Vector generates transfection results in as little as 5 h, while wild-type GFP requires 24 to 48 h. Users can assess eukaryotic cell transfections in 'real time', allowing detection and selection of transfected cells by fluorescence microscopy or flow cytometry.

Reader Enquiry No. 122

Four new **neurotrophic factors** extend the range of growth factors now available from Genzyme Diagnostics. These include recombinant human brain-derived neurotrophic factor, recombinant human NT-3 and NT-4, and recombinant human glial-derived neurotrophic factor. A new recombinant rat ciliary neurotrophic factor is also available. All growth factors are produced from genes expressed in *Escherichia coli* and purified by cation-exchange chromatography and/or reverse-phase HPLC to greater than 95 per cent purity.

Reader Enquiry No. 123

These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.



Green Lantern-1 Vector.

ADVERTISEMENTS

NEWCASTLE UNIVERSITY
MOLECULAR BIOLOGY UNIT

PROTEIN SEQUENCING
Membrane-bound, solid or solution samples

£12 per cycle
5 cycles minimum - no set up charge

For further information contact Dr. Joe Gray
Phone: +44 191 222 8612
Fax: +44 191 222 8129
Email: JOE.GRAY@newcastle.ac.uk

READER ENQUIRY NO 12

DNA/RNA Custom Synthesis
of high quality and purity with high and low scale for all applications

- * HPLC purification, mixed base sites, ANTISENSE, and all oligo modifications available
- * Special Offers Through 8-31-96 and Guaranteed Low Prices

GML
Genetic Medicine Laboratories
Tel: +49-(0)9360-1661

Bahnstrasse 2 D-97450 Arnheim-Brünnfeld
A NEW FRONTIER IN GENE TECHNOLOGY
Fax Ordering: +49-(0)9360-1809

READER ENQUIRY NO 4

DNA SEQUENCING

£45 per reaction

- PCR products
- single strand DNA
- M13 inserts
- double strand DNA
- plasmid inserts
- no result? no charge!

By research labs for research labs

Molecular Biology Services
King's College School of Medicine and Dentistry
Rayne Institute, 123 Goldharbour Lane, London SE5 9NU
Tel: 0171-346 3126 Fax: 0171-733 3577

READER ENQUIRY NO 11